

PEDRO RODRIGUES TORRES JÚNIOR

CARACTERIZAÇÃO DA REDE DE SINCRONIZAÇÃO NA INTERNET

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Informática no Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, Universidade Federal do Paraná.

Orientadora: Prof. Cristina Duarte Murta

CURITIBA

2007

Dedicatória

Dedico este trabalho à minha namorada Renata pela sua paciência em entender quanto esse período foi importante para o meu crescimento pessoal e profissional. Seu apoio incondicional foi fundamental para me dar tranquilidade e motivação.

Agradecimentos

Agradeço aos meus pais e irmãos pelo exemplo de vida e por todo carinho, dedicação e incentivo que me deram.

Agradeço à Professora Cristina Duarte Murta pela orientação deste trabalho, com todo o esforço que empreendeu para a conclusão deste projeto. O tempo que passamos discutindo as fases deste trabalho me empolgaram a dar o melhor de mim para compensar seu verdadeiro apoio. Sua dedicação em cada pequeno detalhe desse trabalho me ajudou a entender o verdadeiro significado da palavra *orientadora*. Nem mesmo a distância geográfica durante um ano deste trabalho comprometeu sua dedicação. Suas críticas e opiniões são muito valiosas tanto no âmbito profissional quanto pessoal.

Agradeço aos meus amigos por entenderem a falta de atenção que lhes tive. Certamente foram alicerces firmes para que eu pudesse trabalhar, pois sabia que estariam esperando pela minha reaproximação.

Agradeço à Deus por ter me dado a felicidade de compartilhar esses momentos com pessoas que merecem a minha consideração.

SUMÁRIO

LISTA DE FIGURAS	vi
LISTA DE TABELAS	viii
RESUMO	ix
ABSTRACT	x
1 Introdução	1
1.1 Caracterização da Rede de Sincronização do NTP	3
1.2 Organização deste Trabalho	5
2 Conceitos e Trabalhos Relacionados	6
2.1 Conceitos em Medição e Sincronização de Relógios	6
2.2 O Protocolo NTP	8
2.2.1 Arquitetura do NTP	10
2.2.2 Troca de Mensagens e Cálculo do <i>Offset</i>	11
2.2.3 O Algoritmo <i>Clock Filter</i>	14
2.2.4 O Algoritmo <i>Clock Selection</i>	15
2.2.5 O Algoritmo <i>Clustering</i>	17
2.2.6 O Algoritmo <i>Clock Combining</i>	18
2.2.7 Variáveis do NTP	19
2.3 Estudos Anteriores sobre o NTP	20
2.4 Propostas de Melhoramento para o NTP	22

2.5	Redes Grandes e Complexas	24
3	Coleta e Organização dos Dados	26
3.1	O <i>Spider</i> NTP	27
3.2	A Primeira Coleta de Dados	27
3.3	A Segunda Coleta de Dados	28
3.4	Organização dos Dados	29
3.5	Questões de Segurança	30
4	Caracterização da Qualidade da Sincronização	32
4.1	Tamanho da Rede de Sincronização	34
4.2	Caracterização da Hierarquia da Rede de Sincronização	35
4.3	Caracterização da Qualidade da Sincronização	37
4.3.1	Análise do Offset	37
4.3.2	Atraso para o Par de Sincronização e a Raiz	40
4.3.3	Atraso para o Par de Sincronização por <i>stratum</i>	43
4.3.4	Dispersão para a Raiz e para o Par de Sincronização	43
4.3.5	Análise do Jitter	45
4.3.6	Estabilidade dos Relógios	48
4.4	Correlação das Variáveis NTP	49
4.5	Síntese do Capítulo	50
5	Aspectos Topológicos da Rede NTP	53
5.1	Visualização da Rede NTP	53
5.2	Localização Geográfica dos Nós	55
5.3	Identificação dos Sistemas Autônomos	63
5.4	Outras Características Topológicas	65
5.5	Síntese do Capítulo	69
6	Conclusão e Trabalhos Futuros	71
	REFERÊNCIAS BIBLIOGRÁFICAS	73

Lista de Figuras

2.1	Grafo da sub-rede de sincronização (esquerda) e árvore de sincronização (direita) do NTP	10
2.2	Organização geral do NTP	12
2.3	Medindo RTT e Offset. As setas indicam o sentido da troca de mensagens	12
2.4	Exemplo de cálculo de RTT e <i>offset</i>	13
2.5	Exemplo do algoritmo <i>clock selection</i>	16
2.6	Exemplo do algoritmo <i>clustering</i>	17
2.7	Métricas do NTP	20
4.1	PDF, CDF e CCDF do <i>offset</i> em relação aos pares e aos servidores.	39
4.2	PDF, CDF e CCDF do RTT para o par de sincronização e para a raiz.	42
4.3	PDF, CDF e CCDF da dispersão para o par de sincronização e para a raiz.	45
4.4	PDF, CDF e CCDF do <i>jitter</i> em relação aos pares e aos servidores	47
4.5	PDF, CDF e CCDF da estabilidade dos relógios	49
4.6	Correlação das variáveis de sistema: atraso para a raiz e <i>offset</i>	50
4.7	Correlação das variáveis de sistema: estabilidade e dispersão para a raiz.	51
5.1	Visualização das associações da rede NTP entre os <i>hosts stratum 2</i> e seus pares de sincronização <i>stratum 1</i>	54
5.2	Visualização das associações da rede NTP entre os <i>hosts stratum 3</i> e seus pares de sincronização <i>stratum 2</i>	56
5.3	Visualização das associações pares de sincronização da rede NTP entre os <i>hosts</i> do nível <i>stratum 3</i> até o nível <i>stratum 16</i>	57

5.4	Distribuição das associações por <i>stratum</i> (esquerda) e dos clientes por <i>stratum</i> (direita).	68
-----	---	----

Lista de Tabelas

2.1	Descrição das principais variáveis do NTP.	19
3.1	Campos da tabela <i>server</i>	30
3.2	Campos da tabela <i>peer</i>	31
4.1	Crescimento da rede NTP ao longo do tempo.	33
4.2	Códigos de origem de sincronização da variável de sistema <i>status</i>	33
4.3	Códigos para a variável <i>status</i> dos pares	34
4.4	Distribuição dos servidores por <i>stratum</i>	36
4.5	Evolução da distribuição (%) de servidores por <i>stratum</i> ao longo do tempo.	37
4.6	Estatísticas para a variável <i>offset</i> do sistema e do par	38
4.7	Evolução da métrica <i>offset</i> ao longo do tempo.	40
4.8	Estatística do RTT para o par de sincronização e para a raiz.	41
4.9	RTT (ms) para o par de sincronização em diversos estudos.	42
4.10	Distância para a raiz (tempo em ms) em diversos estudos.	43
4.11	RTT (ms) para o par de sincronização por <i>stratum</i> em diversos estudos.	43
4.12	Estatística da dispersão para o par de sincronização e para a raiz.	44
4.13	Estatística do <i>jitter</i> da variável de sistema e da variável do par	47
4.14	Estatística da estabilidade dos relógios	48
5.1	Localização dos <i>hosts</i> NTP no mundo.	59
5.2	Localização dos servidores de tempo no mundo.	60
5.3	Distribuição dos pares de sincronização cliente/servidor entre os países (lista parcial).	61

5.4	Países servidores de tempo e número de clientes (lista parcial).	62
5.5	Países com <i>hosts</i> clientes do Brasil.	62
5.6	Quantidade de <i>Hosts</i> NTP em cada sistema autônomo (lista parcial). . . .	64
5.7	Número de pares de sincronização em cada sistema autônomo (lista parcial). .	66
5.8	Relação cliente/servidor entre os sistemas autônomos (lista parcial). . . .	67
5.9	Sistemas autônomos com pares de sincronização no Brasil (lista parcial). .	67
5.10	Índice de carga por <i>stratum</i> em vários estudos.	69

Resumo

Manter a sincronização de relógios é uma tarefa importante e complexa em sistemas distribuídos. O *Network Time Protocol* - NTP - é um protocolo criado para obter e manter a sincronização de relógios de computadores com o tempo real mundial. Para isto, o NTP constrói uma rede lógica dinâmica, com características de uma rede *peer-to-peer*, que se auto-organiza, a partir da inserção inicial dos computadores, por meio de troca freqüente de mensagens entre os pares. O NTP utiliza a Internet como meio de comunicação entre os nodos da rede NTP e é o protocolo padrão para sincronização de relógios de computadores conectados à Internet.

O objetivo deste trabalho é caracterizar, sob vários aspectos, a rede de sincronização de relógios criada pelo NTP na Internet. Com dados coletados em 2005 por um robô, este trabalho apresenta uma caracterização da rede de sincronização descoberta, composta por milhares de nodos, e distribuída geograficamente por mais de 150 países e regiões do planeta. Muitos aspectos que definem a qualidade da sincronização obtida e as características topológicas da rede são apresentados e analisados.

Os resultados da análise são comparados com estudos similares realizados nos últimos 15 anos, mostrando a evolução da rede de sincronização do NTP neste período. Os resultados apresentam evidências do crescimento da rede, da evolução da qualidade da sincronização e da melhoria da infra-estrutura da Internet. No entanto, alguns pressupostos do protocolo NTP podem não ser completamente atendidos na prática, o que pode prejudicar a qualidade da sincronização e limitar os resultados obtidos pelo NTP.

Abstract

Clock synchronization is an important and complex task in distributed systems. The Network Time Protocol - NTP - is a protocol created to build and maintain the synchronization of computer clocks with the world real time. To do so, the NTP implements a dynamic logical network, presenting characteristics of a peer-to-peer network, which keeps itself self-organized, from the initial insertion of the computers, by means of frequent exchange of messages between peers. The NTP uses the Internet as a communication media for network nodes, and it is the standard protocol for synchronization of computer clocks connected to the Internet.

The goal of this work is to characterize, in many aspects, the clock synchronization network created by the NTP in the Internet. With data collected in 2005 by a robot, this work presents a characterization of the NTP network discovered, composed by thousand of nodes, and geographically distributed in over 150 countries and regions of the planet. Many aspects that define the quality of synchronization and timekeeping and the network topological characteristics are presented and evaluated.

The results of this analysis are compared to similar studies carried through in the past 15 years, showing the evolution of the NTP synchronization in this period. The results present evidences of the network growth, the evolution of the synchronization quality and the improvement in Internet infrastructure. However, some assumptions taken by the NTP protocol can be not completely taken in practical situations, harming the synchronization quality and limiting the results obtained with NTP.

Capítulo 1

Introdução

Sincronização de relógios é um aspecto essencial na infra-estrutura de muitos sistemas distribuídos. Inúmeras atividades nas áreas industrial, comercial, financeira, de saúde, transporte, lazer, serviços e outras requerem sincronização para execução de tarefas e registro de tempo de ocorrência de eventos. Tarefas como processamento de transações, escalonamento e controle de processos, autenticação, operações financeiras, processos industriais, diagnóstico de faltas, investigação de crimes e assinatura digital são exemplos de atividades que podem requerer registro de tempo confiável.

No âmbito dos sistemas computacionais, vários programas e aplicações dependem da sincronização correta dos relógios. Por exemplo, sistemas de gerenciamento de arquivos tais como NFS (*Network File System*) e AFS (*Andrew File System*), servidores de banco de dados e de e-mail, sistemas de *back-up*, programas de agendamento de tarefas tais como o `cron` e o `at`, e ferramentas de geração automática de código executável, como o `make`, dependem das estampas de tempo para tomar decisões. Sistemas de caches na Web verificam a validade de um documento com base em sua estampa de tempo, também presente em eventos do SNMP (*Simple Network Management Protocol*). A estampa de tempo em pacotes é fundamental em estudos e medições em redes, bem como em alguns sistemas criptográficos e protocolos de segurança dependentes do tempo, tal como Kerberos [11]. Enfim, cada vez mais processos operando em sistemas distribuídos dependem de registro de tempo confiável. A falta de sincronização pode levar a falha na execução correta das

operações, perda de dados, problemas de segurança e perda de credibilidade de sistemas e empresas. No entanto, cada sistema pode ter necessidades específicas quanto à resolução, precisão, exatidão ou outras propriedades dos relógios do sistema.

Sincronizar relógios é um problema difícil porque o tempo é um alvo móvel. Além disso, observa-se uma necessidade crescente de sincronização com exatidão em escalas de tempo cada vez menores. Muitas tarefas tais como o registro da criação de um arquivo ou da entrada de um funcionário em uma empresa podem ser facilmente executadas com exatidão em uma escala de tempo grande como dia, hora e minuto. No entanto, outras aplicações podem requerer exatidão em escala de segundo ou inferior (*ms*, *μs*, *ns*, *ps*). À medida que os computadores ficam mais rápidos e executam mais tarefas por unidade de tempo, os eventos ocorrem em escalas de tempo cada vez menores, e os registros de estampa de tempo requerem exatidão e precisão cada vez maiores.

Para manter o sincronismo dos relógios diversos protocolos foram propostos, como o *Daytime Protocol* [36], *Time Protocol* [37], *ICMP Timestamp Message* [33] *IP Timestamp Option* [32] e o *Network Time Protocol* – NTP [22]. O *Network Time Protocol* é o protocolo padrão para sincronização de relógios de computadores conectados à Internet. O NTP é um protocolo dinâmico e flexível, capaz de construir e manter uma rede de sincronização com características de uma rede *peer-to-peer*, o que torna o processo confiável e robusto mesmo quando executado em uma rede como a Internet. O principal produto do NTP é a estampa de tempo gerada em cada sistema que executa o protocolo.

Computadores de propósito geral estão se tornando mais rápidos a cada ano, com seus *clocks* operando atualmente em frequências de 3GHz ou mais. Redes de computadores, como a Internet, também estão mais rápidas, com a disseminação de tecnologias para roteamento IP que operam comutando pacotes em velocidades de gigabits por segundo. Isto permite que os pacotes IP experimentem filas menores nos roteadores, diminuindo o atraso e a perda de pacotes.

O avanço na infra-estrutura da Internet também deve representar um avanço na qualidade de sincronização do NTP. Além do melhoramento no *hardware* da infra-estrutura das redes de computadores, melhorias no *software*, como alterações realizadas diretamente

nos sistemas operacionais dos computadores, também contribuem para a qualidade da sincronização de computadores em rede. Por outro lado, o aumento na utilização de enlaces com velocidades assimétricas, como acontece em acessos feitos com tecnologia xDSL, pode prejudicar a sincronização feita pelo NTP. Em geral, neste tipo de rede o atraso percebido em cada sentido da conexão é diferente, o que interfere negativamente no processo de sincronização de tempo realizado pelo NTP.

O protocolo NTP constrói uma rede de sincronização de relógios de máquinas conectadas à Internet, cujo objetivo é sincronizar os relógios destas máquinas com o tempo real mundial. Um cliente NTP é inicialmente configurado pelo administrador com uma lista de pares que fornecem o tempo. A partir da configuração inicial, o protocolo determina, por meio de vários algoritmos, o servidor de tempo mais adequado dentre os disponíveis na rede de sincronização alcançada pelo cliente. Assim, a rede NTP construída é *semi-auto-organizada*, complexa e extremamente dinâmica. Além dos aspectos relacionados à qualidade da sincronização, o estudo da topologia da rede formada também revela informações importantes sobre a infra-estrutura da Internet.

1.1 Caracterização da Rede de Sincronização do NTP

O objetivo deste trabalho é avaliar a qualidade da sincronização dos relógios de computadores conectados à Internet, feita pelo NTP. Vários aspectos que determinam a qualidade da sincronização e da manutenção do tempo na rede são abordados e discutidos, por exemplo, as variáveis que influenciam os algoritmos utilizados pelo protocolo NTP, a topologia da rede de sincronização e a infra-estrutura da Internet. A importância de cada aspecto e variável no processo de sincronização é discutida e avaliada.

Algumas caracterizações foram feitas anteriormente, entre 1989 e 1999 [20, 9, 25, 26]. No entanto, nos últimos seis anos, a Internet cresceu e evoluiu em vários aspectos, desde o número de usuários e máquinas, até a infra-estrutura da rede. Assim, consideramos oportuno fazer uma caracterização da rede NTP atual e comparar nossos resultados com os resultados de caracterizações anteriores, discutindo a evolução da rede e da qualidade da sincronização de relógios de computadores na Internet.

Este trabalho é baseado em coleta de informações sobre os *hosts* NTP alcançáveis, de domínio público, na Internet. As informações coletadas permitiram a caracterização da rede de sincronização e de aspectos da infra-estrutura da Internet. Diversas características do protocolo NTP foram observadas e analisadas. As comparações com estudos anteriores possibilitaram avaliar o comportamento e a evolução da rede, permitindo o melhor entendimento da dinâmica da topologia e revelando novas informações sobre a qualidade de sincronização.

Para realizar o estudo foi utilizado um *spider*. Um *spider*, também conhecido como robô, é um programa que realiza consultas a um grande número de *hosts* na Internet para coletar informações. O *spider* NTP percorreu o grafo da rede de sincronização do NTP, pesquisou e armazenou em uma estrutura de dados as informações específicas do protocolo disponíveis em cada *host*.

Os dados coletados em diversas variáveis do protocolo NTP foram caracterizados, permitindo análise do tamanho da rede, da hierarquia dos servidores e da qualidade de sincronização. A topologia da rede de sincronização foi analisada, bem como a distribuição geográfica dos nodos da rede no planeta. Informações sobre a localização dos nodos da rede nos diversos sistemas autônomos também foram coletadas e analisadas. Alguns grafos da sub-rede de sincronização foram plotados para permitir sua visualização. Além disso, aspectos topológicos que indicam um modelo para a rede de sincronização foram observados.

Os resultados da caracterização indicam crescimento da rede de sincronização ao longo do período, melhoria na qualidade da sincronização, com tempos mais exatos, e melhoria na latência dos caminhos na Internet. A melhoria da qualidade da sincronização pode ser atribuída à diminuição da latência na Internet, uma vez que os erros na medição e propagação do tempo são proporcionais à latência da rede. Observou-se também uma diminuição na porcentagem de servidores que são fontes de tempo para toda a rede, bem como um crescimento na quantidade de computadores que solicitam sincronização mais exata. A análise da topologia da rede revela que o número de associações com pares de sincronização é pequeno, o que pode prejudicar a qualidade da sincronização devido às

características dos algoritmos do NTP. Os algoritmos do NTP escolhem a melhor fonte de tempo a partir de um conjunto maior de fontes de tempo disponíveis. Com poucas associações e poucas fontes de tempo, o número de computadores clientes para cada fonte de tempo é muito grande, o que pode comprometer a disseminação da informação de tempo pela rede. Os dados revelam que a carga nos servidores de tempo primários é alta. A grande variação na distribuição das conexões dos nodos evidencia características do modelo livre de escala, que é utilizado atualmente para modelar redes complexas.

A distribuição geográfica dos nodos da rede pelo mundo e a análise topológica em relação aos sistemas autônomos da Internet revelam que a rede NTP está distribuída em mais de 150 países e regiões, com maior presença em locais economicamente mais desenvolvidos. Os países que possuem mais nodos da rede e mais servidores origem de tempo são os mais ricos. De forma geral, os computadores da rede NTP procuram sincronizar seus relógios com nodos próximos, no mesmo sistema autônomo ou no mesmo país.

Este trabalho contribui para o estudo empírico da sincronização de relógios em sistemas distribuídos, particularmente na Internet. O trabalho apresenta dados que revelam a situação atual da rede NTP, de forma qualitativa e quantitativa. Os resultados podem ser empregados em sistemas distribuídos construídos sobre a Internet, em especial, grades computacionais e *clusters*. Além disso, os resultados podem ser úteis para a proposição de melhorias no protocolo NTP ou para a proposição de novos protocolos para sincronização de relógios.

1.2 Organização deste Trabalho

Este trabalho está organizado como descrito a seguir. No capítulo 2 são apresentados conceitos no contexto de sincronização de relógios, aspectos do protocolo NTP essenciais ao entendimento desse trabalho e um resumo dos trabalhos relacionados. O capítulo 3 descreve como foi feita a coleta e a organização dos dados. O capítulo 4 descreve os resultados da caracterização da qualidade da rede de sincronização do NTP. O capítulo 5 descreve os resultados da análise de vários aspectos da topologia da rede de sincronização. Por fim, o capítulo 6 apresenta as conclusões e as possibilidades de trabalhos futuros identificadas no desenvolvimento desta pesquisa.

Capítulo 2

Conceitos e Trabalhos Relacionados

Este capítulo apresenta os principais conceitos relativos à sincronização de relógios em sistemas distribuídos, bem como diversos trabalhos que descrevem o estado da arte das pesquisas relacionadas ao tema.

A organização deste capítulo é feita em cinco seções. A Seção 2.1 apresenta alguns conceitos fundamentais relacionados ao tempo em sistemas distribuídos; a Seção 2.2 discorre sobre as principais características do protocolo NTP; a Seção 2.3 apresenta os resultados de estudos anteriores sobre a rede NTP; a Seção 2.4 cita algumas propostas de melhoria para o protocolo NTP. Por fim, a Seção 2.5 apresenta alguns conceitos relacionados com redes complexas.

2.1 Conceitos em Medição e Sincronização de Relógios

Uma forma comum de medir o tempo é utilizando a transição do sol. Esta medição astronômica está relacionada com a rotina de cada dia, com períodos bem definidos, como a noite e o dia, que se repetem. Um dia solar é o tempo que o sol leva para assumir a mesma posição no céu após iniciada uma medição. A unidade padrão de tempo, o segundo, é definida como a fração $1/86400$ do dia solar. Outra forma de medir o tempo é através da medição atômica. A medição atômica só foi possível ser feita com a invenção do primeiro relógio atômico em 1948. Esta forma de medição define o segundo como o tempo necessário para que o átomo de césio 133 realize 9.192.631.670 transições [4].

O período solar tende a divergir em relação ao período atômico. As voltas da terra em torno do sol não gastam sempre o mesmo tempo atômico, ou seja, um dia solar não contém a mesma quantidade de segundos de um relógio atômico. Em geral, isso ocorre pois a terra está se afastando do sol, e devido a outros fenômenos astronômicos. A medição atômica é definida como o Tempo Atômico Internacional — TAI.

O padrão *Coordinated Universal Time* — UTC — é um padrão internacional baseado no período atômico, mas disciplinado com o período solar. Para fazer a correção do padrão UTC, é utilizado um fator de ajuste de 1 segundo, chamado *leap second*, o qual é inserido ou removido a cada 18 meses aproximadamente. Sinais UTC são sincronizados e difundidos por estações de rádio na terra e através de satélites, para quase todas as partes do mundo. As técnicas existentes para sincronização de relógios em sistemas distribuídos utilizam esses sinais como referência.

Apesar da dificuldade para definirmos o *tempo*, iremos considerá-lo como o intervalo ou período entre eventos ou o momento indicado por um dispositivo de medição, o *relógio*. Um relógio consiste de três mecanismos: um oscilador, um contador de oscilações que registra continuamente o número de ciclos do oscilador, e um mecanismo para mostrar, gravar ou exibir o resultado. Um oscilador é um mecanismo que repete eventos a uma taxa constante, cujo erro é normalmente expresso em PPM (*partes-por-milhão*). O contador converte o número de ciclos contados no oscilador em unidades de tempo conhecidas, tais como segundos, minutos, horas. Cada resultado, ou seja, cada valor do contador é uma estampa do tempo (*timestamp*).

Sincronização é a coordenação de eventos em relação ao tempo. Há dois propósitos para a sincronização. O primeiro propósito é permitir e assegurar que eventos ocorram em uma determinada seqüência, ou em um tempo pré-definido. Neste caso a sincronização é necessária para disparar eventos e registrar sua ocorrência para que, posteriormente, seja possível fornecer prova de quando os eventos ocorreram e em que seqüência. O segundo propósito é recuperar a informação de tempo registrada, relativa a eventos ocorridos, sempre que necessário. Esta tarefa é possível apenas se estampas de tempo exatas e precisas estão disponíveis. Por exemplo, para analisar certos fatos e ocorrências, muitas vezes é

necessário ter uma sequência de eventos organizada em ordem cronológica e sincronizada com o tempo aceito mundialmente.

Vários conceitos são relacionados à sincronização de relógios e manutenção de tempo. Exatidão (*accuracy*) é o grau de conformidade de uma medida de tempo em relação ao valor verdadeiro ou referência padrão, neste caso conhecida como a “hora certa” (UTC). Resolução ou precisão é a unidade da menor estampa de tempo que um relógio pode gerar, seja minuto, segundo, ou unidades menores. A diferença de tempo entre dois relógios é denominada *time offset* e a diferença de frequência é denominada *frequency offset*. A estabilidade de um relógio é uma medida das flutuações de frequência do oscilador. As instabilidades no oscilador, chamadas de *drift*, podem ser causadas por radiação, pressão, temperatura e umidade. Mudanças internas no oscilador durante sua vida útil também causam instabilidades.

Uma das propriedades mais importantes de um relógio é a monotonicidade, que significa que o tempo é sempre crescente, ou seja, duas leituras consecutivas de um relógio devem prover tempo maior para a segunda leitura do que o tempo fornecido na primeira leitura, ou no máximo igual, dependendo da resolução. Relógios baseados em osciladores precisos sempre mantêm esta propriedade. No entanto, relógios implementados em software podem facilmente ter sua estampa de tempo alterada para um tempo no passado. Muitas aplicações e protocolos podem não ser programados para esta situação, causando erros de software, o que torna a monotonicidade uma propriedade extremamente importante. Em várias situações, a monotonicidade é mais importante do que a exatidão.

2.2 O Protocolo NTP

O *Network Time Protocol* — NTP — [22] é um protocolo para sincronização de relógios de sistemas computacionais que opera sobre uma rede comutada de pacotes.

O NTP é um dos protocolos mais antigos em uso na Internet [23] e foi desenvolvido para manter precisão e robustez, mesmo quando usado sobre caminhos que envolvam vários roteadores e redes não confiáveis, que causam variação no atraso e na alcançabilidade no destino de uma mensagem enviada, devido à carga da rede, seleção de rotas e outros

eventos, como é o caso da Internet. A arquitetura do NTP provê o serviço de tempo UTC, compensando, através de técnicas estatísticas, as instabilidades observadas nas mensagens trocadas sobre a rede IP. O NTP foi projetado para prover escalabilidade, permitindo que todos os computadores da Internet possam se manter sincronizados através de uma grande rede de sincronização. A descrição do protocolo feita nessa seção refere-se à versão 4, descrita em [13]. O desenvolvimento da versão 4 foi iniciado em 1999 e atualmente é descrito em um *draft* do *Internet Engineering Task Force* (IETF).

O NTP define o conceito de *stratum*, que é uma classificação dada aos nodos da rede, com o objetivo de dirigir a sincronização em múltiplos níveis. Os nodos são classificados em *stratum* 1 a 16, e nodos de *stratum* i fornecem tempo para nodos cujo *stratum* é $i + 1$. *Servidor* é um nodo que fornece a informação de tempo para fins de sincronização, enquanto *cliente* é o nodo que recebe essa informação de tempo. Um nodo cliente também pode operar simultaneamente com a função de servidor, permitindo a sincronização sucessiva até o mais baixo nível de *stratum*. Nesse trabalho chamaremos de *host NTP*, ou simplesmente *host*, qualquer nodo atuante na rede do NTP; e chamaremos de *par NTP*, ou simplesmente *par*, qualquer referência a um relógio de um *host NTP* remoto, ou mesmo local, através de algum dispositivo que esteja fornecendo o tempo, como um relógio atômico conectado via interface serial.

A rede lógica de nodos NTP é conhecida como *sub-rede de sincronização* e pode ser representada no formato de um grafo [3]. Na raiz de uma sub-rede de sincronização estão os *hosts* primários, chamados *stratum 1*, que se conectam diretamente a uma fonte confiável de tempo como, por exemplo, um rádio, um receptor de sinal de satélite ou um modem telefônico recebendo informação de tempo UTC. *Hosts* secundários, chamados *stratum 2*, são sincronizados com servidores *stratum 1*, e assim sucessivamente até os *hosts* ditos *stratum 16*.

Cada cliente na sub-rede de sincronização é configurado com um nodo atuando como servidor ou *par de sincronização*. Para permitir redundância e confiabilidade, mais de uma opção para par de sincronização é desejável. Uma *associação* é uma referência a um nodo capaz de atuar como par de sincronização e pode estar explícita na configuração

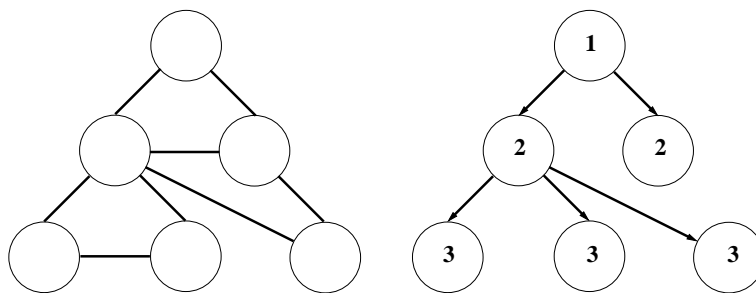


Figura 2.1: Grafo da sub-rede de sincronização (esquerda) e árvore de sincronização (direita) do NTP

ou ainda ser criada pelo *daemon* NTP em tempo real. No exemplo da Figura 2.1 temos à esquerda uma sub-rede de sincronização do NTP em que os nodos são *hosts* NTP e as linhas representam associações. Alguns nodos possuem diversas associações. No lado direito da Figura 2.1, algumas associações foram eleitas pelo protocolo para fornecer o tempo UTC, formando uma árvore de sincronização. As linhas indicam a direção do fornecimento de tempo entre o par de sincronização e o cliente. Na raiz da árvore temos um nodo *stratum* 1, indicando ser um servidor que consulta diretamente uma fonte de tempo confiável. Os nodos *stratum* 2 representam um servidor de segundo nível da árvore, que utilizam o servidor *stratum* 1 como fonte do tempo e fornecem sincronização para os nodos *stratum* 3.

Diversas métricas são utilizadas pelo protocolo NTP para permitir a escolha da associação que atuará como par de sincronização. A principal métrica é o valor do *stratum*. Com isso o NTP pode optar por utilizar preferencialmente um par de sincronização mais preciso, pois está mais próximo de uma fonte confiável de tempo. Veremos nas próximas seções quais são as outras métricas e como são utilizadas pela máquina de estado do NTP para garantir melhor confiabilidade e qualidade na sincronização do tempo.

2.2.1 Arquitetura do NTP

Para entender o fluxo de processamento do NTP desde as primeiras trocas de mensagens até o ajuste do relógio local, é essencial conhecer os componentes desse sistema. A Figura 2.2 mostra esses componentes. Inicialmente, o processo que implementa o protocolo

NTP no sistema, chamado *daemon NTP*, controla o intervalo de troca de mensagens para cada associação mantida no *host*. Após as primeiras trocas de mensagens é possível calcular o valor do *time offset*, ou simplesmente *offset*, do atraso de ida e volta (*round-trip time* ou RTT), ou simplesmente atraso, além de outras métricas. Esses valores são mantidos em variáveis separadas, relacionadas com cada associação, e são analisados pelo algoritmo chamado *clock filter*, o qual selecionará a melhor amostra dentre as últimas calculadas.

As amostras selecionadas seguirão para o processo de seleção executado pelo algoritmo *clock selection*. Amostras de associações que indicam erros ou inconsistências, em comparação aos demais valores, são descartados nessa fase. Essas associações são denominadas *false-tickers*. Apenas associações que mantêm exatidão com um padrão confiável de tempo passarão para a próxima fase. Essas associações são denominadas *true-chimers*.

A seguir, o algoritmo de *clustering* selecionará, dentre as *true-chimers*, os valores mais confiáveis e exatos. Assumindo um número suficiente de *true-chimers*, o algoritmo utiliza diversas estatísticas para escolher os que seguirão para a próxima fase, chamada de *clock combining*. Na fase de *clock combining*, as associações com amostras que chegaram até este ponto, denominadas sobreviventes, determinarão o termo médio dos valores de *offset* (uma combinação), levando principalmente em consideração a métrica chamada *distância de sincronização*, que é a distância estimada pelo protocolo para a raiz da sub-rede de sincronização.

Até este ponto, um valor de ajuste para o relógio local já foi determinado, para colocá-lo em fase com relação ao UTC. Este valor é recalculado a cada recebimento de mensagem que resulte em novo *offset* de atualização do relógio local. Os ajustes no sistema são aplicados pelo algoritmo *disciplina do relógio local*.

As subseções a seguir descrevem cada um dos algoritmos NTP citados acima.

2.2.2 Troca de Mensagens e Cálculo do *Offset*

O NTP utiliza o protocolo *Internet Protocol* — IP [35] e o protocolo *User Datagram Protocol* — UDP [34] para a troca de mensagens. Como o protocolo de transporte utilizado é não confiável e não orientado a conexão, o próprio NTP provê meios para detectar

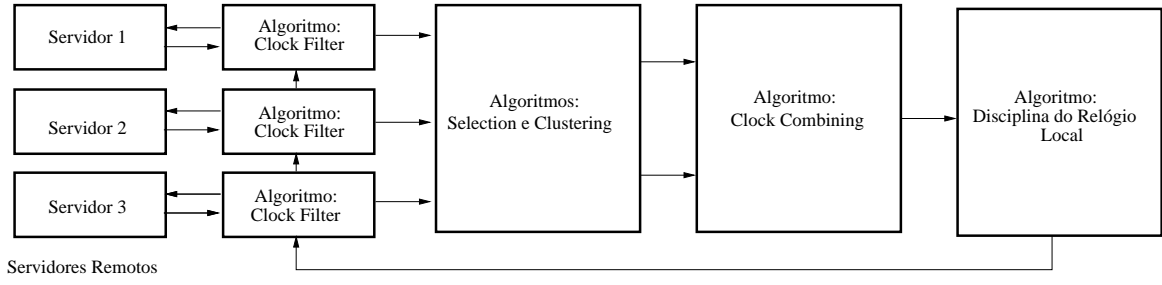


Figura 2.2: Organização geral do NTP

pacotes duplicados, integridade dos dados e também garantir autenticidade da informação e do par [21].

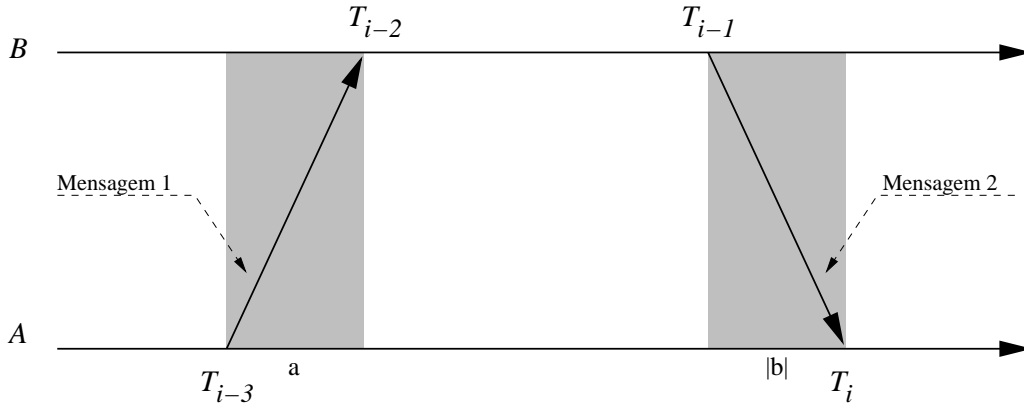


Figura 2.3: Medindo RTT e Offset. As setas indicam o sentido da troca de mensagens

A Figura 2.3 mostra como os *timestamps* NTP são numerados e trocados entre os *hosts* A e B. A operação inicia-se quando A lê o seu relógio, no tempo T_{i-3} , salva esse valor em uma variável e envia a mensagem 1 contendo T_{i-3} para B. Na chegada da mensagem, B lê o tempo T_{i-2} e armazena esse valor e T_{i-3} em variáveis locais. Algum tempo depois, B lê o seu relógio no tempo T_{i-1} e envia a mensagem 2 contendo T_{i-3} , T_{i-2} , T_{i-1} para A. Na chegada da mensagem A lê o tempo T_i . Assim, o *host* A tem quatro *timestamps*, T_{i-3} , T_{i-2} , T_{i-1} e T_i , necessários para calcular o *offset* e o RTT.

Seja $a = T_{i-2} - T_{i-3}$ e $b = T_{i-1} - T_i$. Se a diferença de atraso na rede entre A e B e entre B e A é pequena, ou seja, o caminho é simétrico, então o RTT δ_i e o *offset* θ_i do *host* A em relação ao *host* B no tempo T_i são próximos a:

$$\delta_i = a - b \quad \text{e} \quad \theta_i = \frac{a + b}{2}$$

Desta forma, tanto o *host A* quanto o *host B* podem, independentemente, calcular o RTT e o *offset* usando uma simples troca bidirecional de mensagens. No exemplo da Figura 2.4, considere os relógios dos *hosts A* e *B* estáveis e uma rede com caminhos simétricos. Podemos calcular pela escala na linha de tempo que o *offset* de *A* em relação a *B* é de -3 . Seja $T_{i-3} = 15$, $T_{i-2} = 14$, $T_{i-1} = 17$ e $T_i = 22$, então $a = -1$ e $b = -5$. O valor de RTT é $-1 - (-5) = 4$ e o *offset* é $\frac{-1+(-5)}{2} = -3$.

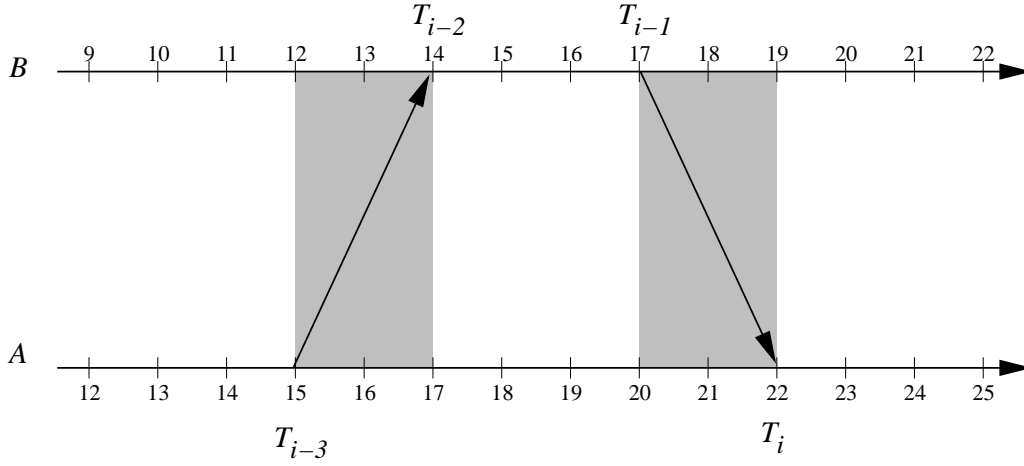


Figura 2.4: Exemplo de cálculo de RTT e *offset*

Tal como apresentado, o cálculo do *offset* e do RTT considera o caminho de ida e o caminho de volta idênticos, não levando em consideração erros estocásticos, observados no atraso em filas de roteadores e *switchs*. A Internet também sofre variação no atraso devido a roteamento assimétrico, enlaces de diferentes velocidades e diversas outras causas. Sendo assim, considere o *offset* verdadeiro de B em relação a A, denominado θ . Seja x o atraso entre a partida de uma mensagem de A e sua chegada em B. Então $x + \theta = T_{i-2} - T_{i-3} \equiv a$. Como x deve ser um número positivo em nosso universo, $x = a - \theta \geq 0$, o que requer que $\theta \leq a$. Uma analogia requer $b \leq \theta$, assim certamente $b \leq \theta \leq a$, que pode ser expresso da seguinte maneira:

$$b = \frac{a + b}{2} - \frac{a - b}{2} \leq \theta \leq \frac{a + b}{2} + \frac{a - b}{2} = a$$

Representando de outra maneira, o *offset* verdadeiro deve estar entre o intervalo de tamanho igual ao atraso em um único sentido medido (metade do RTT) e centralizado no *offset*:

$$\theta_i - \frac{\delta_i}{2} \leq \theta \leq \theta_i + \frac{\delta_i}{2}$$

2.2.3 O Algoritmo *Clock Filter*

Os valores de *delay* e *offset* (δ , θ) são essenciais para o bom funcionamento do protocolo NTP. O algoritmo *clock filter*, mostrado na organização geral do protocolo, apresentada na Figura 2.2, é desenvolvido para selecionar a melhor amostra, rejeitando aquelas que sofreram alguma interferência devido a colisão de pacotes ou congestionamento na rede. O algoritmo *clock filter* foi proposto observando-se o comportamento de redes de pacotes comutados, que freqüentemente operam abaixo da vazão máxima e, dessa forma, as filas são pequenas e com apenas algumas rajadas. Além disso, os protocolos de roteamento operam minimizando o número de saltos, portanto, diminuindo também o número de filas. Assim, a probabilidade de uma mensagem NTP experimentar uma fila congestionada em uma direção é baixa, bem como a probabilidade de uma simples troca de mensagens encontrar fila congestionada nas duas direções é menor ainda. Desta forma, a melhor amostra de *offset* deve ocorrer com um pequeno atraso.

Esta observação sugere o uso do *minimum filter*. Este filtro utiliza os valores de *offset* θ_0 e RTT δ_0 como calculados no exemplo da Figura 2.4. Além disso, é utilizada uma métrica de qualidade chamada dispersão, ε_0 , que é atribuída com o erro estimado do relógio do servidor remoto, informada junto com a mensagem NTP e, finalmente, o valor do *timestamp* T_i .

O *minimum filter* é composto de um registrador de deslocamento que possui n amostras $(\theta_i, \delta_i, \varepsilon_i, t_i)$, $0 \leq i < n$. Cada nova amostra calculada na troca de mensagens $(\theta_0, \delta_0, \varepsilon_0, t_0)$ substitui a entrada mais antiga. O conteúdo do registrador é copiado para uma lista temporária que é ordenada por uma métrica que utiliza o valor de atraso calculado para cada amostra, e também o tempo de existência da amostra. As amostras antigas têm sua

atuação influenciada pela estabilidade do oscilador do relógio do par, conforme discutido em [39]. Esse fenômeno é chamado de compromisso de *Allan intercept*. A métrica de ordenação prioriza amostras relacionadas com um atraso menor e mais recentes.

A primeira entrada $(\theta_0, \delta_0, \varepsilon_0, t_0)$ na lista ordenada representa a amostra com o menor atraso, que é a utilizada para atribuir a variável *offset* do par $\theta = \theta_0$ e a variável RTT do par $\delta = \delta_0$. A dispersão para o par é calculada utilizando as amostras da lista temporária como segue:

$$\varepsilon = \sum_{k=0}^{n-1} \frac{\varepsilon_k}{2^k + 1}$$

Finalmente, a lista temporária é podada de acordo com o compromisso de *Allan intercept*, deixando m amostras ($0 \leq m < n$). A métrica *jitter* φ , que é utilizada pelo algoritmo de *clustering* como uma métrica de qualidade, é calculada como segue:

$$\varphi = \left(\frac{1}{m-1} \sum_{k=1}^{m-1} (\theta_k - \theta_0)^2 \right)^{\frac{1}{2}}$$

A métrica *jitter* φ é equivalente ao desvio padrão dos *offset* para m amostras, considerando θ_0 como o valor verdadeiro.

2.2.4 O Algoritmo *Clock Selection*

Para garantir uma sincronização confiável, é importante que o NTP utilize servidores redundantes e com diversos caminhos na rede quando possível. Dentre esses pares, é preciso descobrir quais estão com algum erro no tempo fornecido (*falsetickers*) e quais são confiáveis (*truechimers*). O algoritmo *clock selection* determina, dentre todas as associações, um subconjunto com *truechimers* capaz de prover o tempo mais exato e confiável, usando princípios similares aos utilizados em [42].

Conforme apresentado na Seção 2.2.2, o verdadeiro *offset* é aquele que está operando em um intervalo chamado *intervalo de confiança*, que para o NTP, é delimitado pela *distância para a raiz*. As variáveis mantidas pelo NTP para a raiz representam estatísticas acumuladas do caminho completo até o *host stratum 1*. A distância para raiz ou distância

de sincronização do *host* i é definida como:

$$\Lambda_i = \frac{\Delta_i}{2} + E_i$$

onde Δ_i e E_i são o atraso e a dispersão, acumulados até a raiz, respectivamente.

Para encontrar o melhor subconjunto de *truechimers*, uma derivação do algoritmo definido em [12] é utilizada. O algoritmo busca o menor *intervalo de interseção* contendo o intervalo de confiança dos m relógios, dos quais não mais que f são considerados incorretos. O intervalo de interseção contém $m - f$ pontos do intervalo de confiança e pelo menos $m - f$ pontos centralizados nos intervalos de confiança.

A Figura 2.5 mostra a operação do algoritmo em um cenário envolvendo quatro relógios A , B , C e D , onde o intervalo de confiança para cada relógio está centralizado no valor de *offset* e delimitado pela distância de sincronização, porém qualquer ponto nesse intervalo pode representar o verdadeiro *offset* associado com o relógio. Se todos os relógios estão corretos, deve existir um intervalo de interseção não vazio incluindo todos os quatro intervalos de confiança, o que não é o caso da figura. O relógio D , um *falseticker*, é excluído do intervalo de interseção, que inclui os outros três relógios.

Como o *offset* associado a cada relógio está contido no intervalo determinado pelo algoritmo *clock selection*, uma média ponderada desses *offsets*, que será calculada pelo algoritmo *clock combining*, também deve estar contida nesse intervalo.

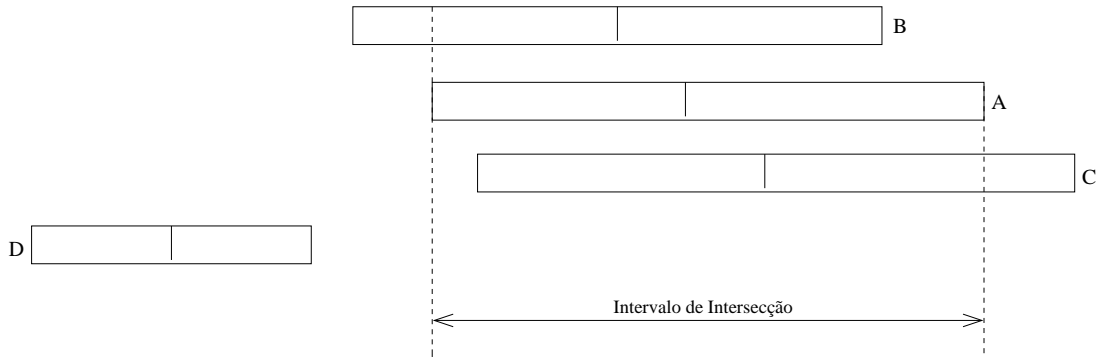


Figura 2.5: Exemplo do algoritmo *clock selection*

2.2.5 O Algoritmo *Clustering*

Após a seleção dos *truechimers* pelo NTP, a partir de uma determinada quantidade de pares, o algoritmo de *clustering* identifica novos sobreviventes para garantir a melhor exatidão. Os *truechimers* selecionados provavelmente não apresentam o mesmo *offset*, mesmo se cada um estiver sincronizado próximo a uma fonte confiável de tempo, por uma razão ou outra. Isso ocorre em geral devido a erros estocásticos na rede comutada de pacotes (Internet), onde o NTP está operando. Os critérios de filtragem utilizados pelo algoritmo *clustering* estão relacionados a uma associação que tenha um baixo valor de *stratum* e uma pequena distância para a raiz. Além disso, o *jitter* calculado pelo algoritmo *clock filter* para a associação é utilizado como métrica de qualidade na escolha dos sobreviventes.

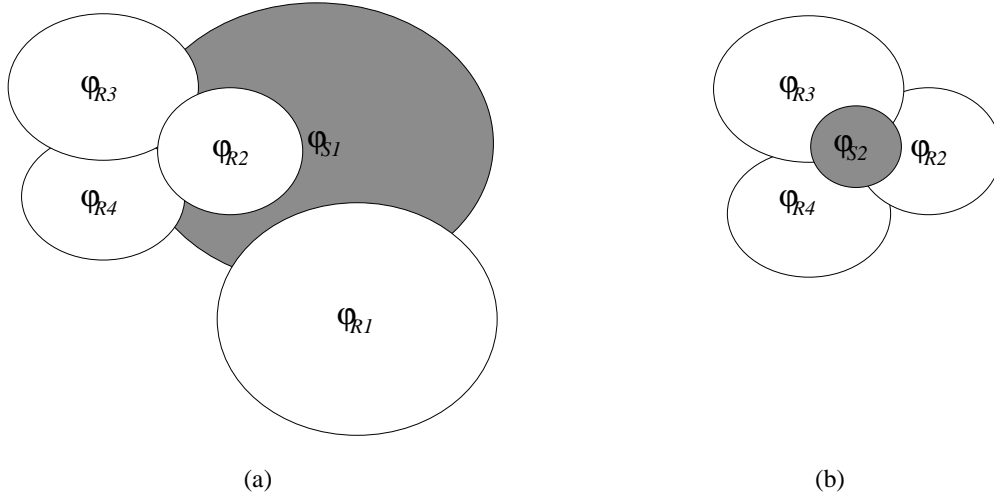


Figura 2.6: Exemplo do algoritmo *clustering*

A Figura 2.6 ilustra o funcionamento do algoritmo. Nesse exemplo, o processo inicia com quatro sobreviventes à esquerda, onde os círculos em branco representam a variação do *offset* (*jitter*) para o par e o círculo cinza representa o maior *jitter* selecionado. O *jitter* selecionado, por sua vez, é calculado para cada par i como segue:

$$\varphi_{s,i} = \left(\sum_{j=0}^{m-1} (\theta_i - \theta_j)^2 \right)^{\frac{1}{2}}$$

onde m é a quantidade de *truechimers* do algoritmo *selection*. Como o maior *jitter* sele-

cionado é maior que o menor *jitter* para o par, o par com o maior produto composto por seu *jitter*, distância para a raiz e *stratum* é removido da lista. No exemplo, *R1* representa o par a ser retirado, deixando os três sobreviventes à direita. Podemos observar que, após o re-cálculo, o maior *jitter selecionado* é menor que o menor *jitter* para o par, e assim o algoritmo termina.

2.2.6 O Algoritmo *Clock Combining*

Os algoritmos descritos anteriormente têm como objetivo selecionar *hosts* baseados em valores baixos de *stratum* e de distância de sincronização. Esses *hosts*, por sua vez, estão conectados a diversos outros, através de diversos caminhos, e podem utilizar diferentes fontes primárias de tempo. Embora esses caminhos não sejam usados diretamente para a sincronização, é possível aumentar a exatidão calculando a média dos *offsets* dos pares sobreviventes, utilizando um peso apropriado para isso.

Se o algoritmo de *clustering* retornou um único sobrevivente, este será o par de sincronização e o algoritmo *clock combining* não será utilizado. Caso contrário, a correção do relógio local será determinada pelo algoritmo *clock combining*. Nos dois casos, o melhor par irá determinar o valor do *stratum*, o atraso e a dispersão acumulada até a raiz, para o *host* NTP.

Sejam θ_i , φ_i e Δ_i o *offset*, o *jitter* e a distância para a raiz para o i -ésimo sobrevivente, respectivamente. Então o *offset* e o *jitter* combinado para o par são, respectivamente:

$$\Theta = a \sum_i \frac{\theta_i}{\Delta_i} \quad \text{e} \quad \varphi_r = (a \sum_i \frac{\varphi_i^2}{\Delta_i})^{\frac{1}{2}}$$

onde a é o normalizador:

$$a = (\sum_i \frac{1}{\Delta_i})^{-1}$$

O resultado Θ será o *offset* utilizado no algoritmo de disciplina do relógio. O *jitter* relacionado do sistema é definido como:

Representação	Significado
θ	<i>Offset</i> para o par
Θ	<i>Offset</i> do sistema
δ	Atraso para o par
Δ	Atraso para a raiz
ε	Dispersão para o par
E	Dispersão para a raiz
φ	<i>Jitter</i> para o par
v	<i>Jitter</i> do sistema

Tabela 2.1: Descrição das principais variáveis do NTP.

$$v = (\varphi_r^2 + \varphi_s^2)^{\frac{1}{2}}$$

Esse valor representa a melhor estimativa de erro na computação do *offset* do relógio e é interpretado como a estatística do erro esperado disponível para as aplicações.

2.2.7 Variáveis do NTP

Como foi visto nos algoritmos descritos nas seções anteriores, o NTP mantém diversas métricas e informações de estado em variáveis. A classificação das variáveis é descrita na especificação do protocolo [38]. As *variáveis do sistema* são informações do ambiente local de uma máquina, relativas ao sistema operacional e ao relógio local. As *variáveis dos pares* representam o estado do protocolo em cada associação. *Variáveis dos pacotes* são aquelas que fazem parte das mensagens do NTP. E, finalmente, os *parâmetros* são constantes utilizadas nas implementações do protocolo. As mesmas variáveis podem ser comuns a duas ou mais classes. Em particular, variáveis de origens distintas podem ter o mesmo nome. A Tabela 2.1 sumariza as principais variáveis do NTP.

A Figura 2.7 apresenta as principais métricas determinadas ou calculadas pelo par ou pelo sistema na arquitetura do NTP. As variáveis dos pacotes englobam, dentre outras variáveis, os quatro *timestamps* obtidos da troca de mensagens (T_{i-3} , T_{i-2} , T_{i-1} e T_i), o atraso até a raiz Δ_R e a dispersão até a raiz E_R herdada do caminho entre o servidor R e a fonte primária de tempo. As variáveis dos pares são calculadas usando as variáveis

de pacotes e incluem o *offset* θ , o RTT δ , e a dispersão ε . O *jitter* φ , que é calculado utilizando as amostras de *offset*, também é uma variável dos pares. As variáveis do sistema são as calculadas pelos algoritmos *selection*, *clustering* e *combining*, incluindo o *offset* Θ , o atraso até a raiz Δ , a dispersão até a raiz E e o *jitter* do sistema v .

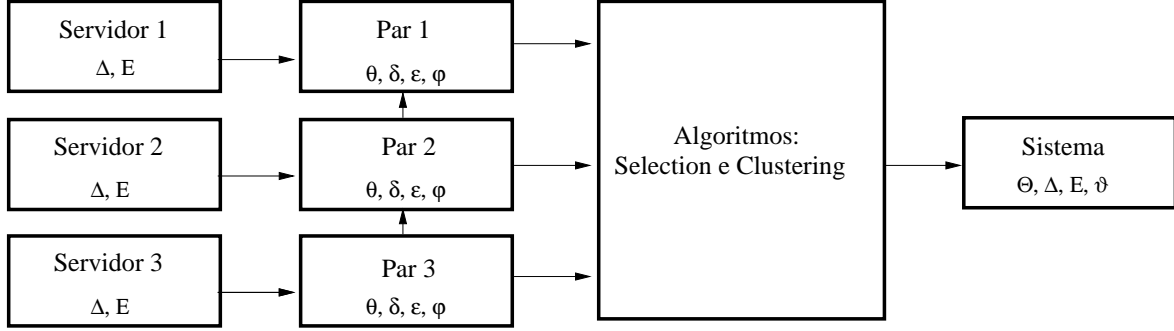


Figura 2.7: Métricas do NTP

2.3 Estudos Anteriores sobre o NTP

Em 1985, David L. Mills fez experimentos na rede de sincronização de tempo [16]. O protocolo avaliado pertencia à família de protocolos locais do *Distributed Computer Network* (DCN) [14]. Apesar de ter sido um teste com apenas 6 *hosts* e 3 relógios de referência, o experimento utilizou diversos tipos de enlaces de comunicação que não operavam exclusivamente para o experimento. Assim, foi possível avaliar o efeito da rede na exatidão dos relógios individualmente. Esse estudo serviu para avaliar o comportamento dos algoritmos utilizados na época e como motivação para o desenvolvimento da primeira versão do *Network Time Protocol*. Ainda no mesmo ano foi feito um outro estudo com 1775 *hosts* e 110 gateways listados em uma tabela de informações do *Network Information Center* (NIC). Destes, 163 responderam para o *Time Protocol* e 504 para o *ICMP Timestamp*. Os resultados de cada protocolo foram analisados e comparados. Ambos apresentaram erros grosseiros gerados por possíveis problemas de implementação [15].

Uma nova coleta e análise de dados foi feita por David L. Mills em 1988 [18] sobre 5498 *hosts* e 224 *gateways* listados no *NIC*. Desses, 1158 responderam para o *Time Protocol* e 1963 para o *ICMP Timestamp*. Desse total, apenas 10% tinham erro menor ou igual a 11

milissegundos.

Após a implementação do *Network Time Protocol* versão 0 [17], da versão 1 [18] e da versão 2 [19], em 1989 David L. Mills conduziu um novo estudo sobre a rede de sincronização [20]. Com o uso mais comum do sistema *Internet Domain Name System* (DNS) para resolver um nome de um *host* para um endereço IP [27], foram utilizados os endereços recém compilados de um estudo sobre servidores de nome, para fazer as consultas, possibilitando encontrar, a partir das consultas iniciais, 107.799 diferentes endereços. Neste experimento foram pesquisadas informações sobre o *Time Protocol*, *ICMP Timestamp* e o NTP. Em resumo, 94.260 endereços resultaram em alguma informação estatística, mas somente 22.758 tinham informações que podiam ser validadas e, desses, apenas 789 responderam ao protocolo NTP. Na comparação feita, o NTP se mostrou mais exato que os outros protocolos.

Em 1993, com a versão 3 do protocolo NTP disponível [22], James Guyton e Michael Schwartz elaboraram uma nova coleta e estudo que buscava *hosts* da rede NTP automaticamente. Esse método, conhecido como *spidering* nos termos Web, iniciou a busca a partir de uma lista de hosts conhecidos que executavam o NTP. Para cada *host*, uma mensagem era enviada solicitando a lista de clientes que tinham feito alguma consulta. A lista de clientes era consultada recursivamente. Esse critério de busca localizou, apesar de restrições enfrentadas, 15.000 hosts. Para cada um desses hosts foram feitas consultas específicas sobre o estado do NTP no sistema local. Sete mil e duzentos *hosts* retornaram dados que permitiram uma análise estatística. Os resultados mostraram que servidores primários NTP estavam sobrecarregados e que a árvore NTP era mal balanceada [9].

Em 1995 David L. Mills realizou um novo estudo sobre a rede de sincronização do NTP [25]. Os servidores NTP foram explorados a partir de uma lista conhecida de servidores públicos com uma técnica similar à utilizada em 1993 [9]. Foram identificados 38.722 *hosts* NTP. O estudo mostrou dados sobre a quantidade de pares que interrogavam esses *hosts*, classificados pelo valor do *stratum*, ou seja, foi medida a carga sobre cada conjunto de servidores, agrupados por *stratum*. Foi identificado um valor máximo de 652 associações para um servidor e uma média de 13,2 associações por servidor. Esse estudo

também mostrou a possibilidade de erros devido à existência de caminhos assimétricos. O resultado mais perturbador encontrado foi a quantidade de servidores *stratum* 1 com parâmetros inadequados ou com configuração errada. Dos 231 servidores primários localizados, 88 apresentavam algum tipo de problema. Em geral, isso se deve ao fato de o NTP depender de configurações fornecidas pelo administrador e que raramente são monitoradas.

Finalmente, em 1999, um novo estudo foi elaborado por Nelson Minar [26]. As informações foram coletadas na árvore de sincronização do NTP também utilizando uma ferramenta de busca de informação automática, como adotado nos estudos anteriores. Foram identificados 175.527 servidores NTP que, comparado com o estudo de dois anos antes, mostrou um crescimento de aproximadamente 3,5 vezes.

Várias comparações com os estudos anteriores demonstraram um comportamento similar da rede NTP, mas algo novo foi observado: do estudo de 1997 para o de 1999 o *offset* diminuiu bastante. A mediana, que era de 20,1 ms, caiu para 1,8 ms, e a média caiu de 28,7 ms para 8,2 ms. Outra métrica que também apresentou uma diferença considerável foi o RTT. A média caiu de 186 ms para 33 ms e a mediana de 118 ms para 32 ms. O estudo indicou que essa diferença se devia a um avanço na estrutura da Internet, mas que também poderia indicar uma mudança na própria topologia da rede de sincronização do NTP.

Um resultado surpreendente refere-se à fonte primária de tempo. Do total de 957 servidores *stratum* 1, somente 28% mostrava ser uma boa referência do tempo, com uma grande quantidade de servidores secundários dependendo desses servidores primários. Esse problema indica queda na qualidade de sincronização de todos os pares que utilizam os *hosts* sem exatidão.

2.4 Propostas de Melhoramento para o NTP

Os estudos feitos na rede de sincronização do protocolo NTP evidenciaram alguns problemas tais como escassez de servidores *stratum* 1, o problema do protocolo operando em caminhos assimétricos, o desbalanceamento de carga nos servidores da rede e outros. Mos-

traremos nessa seção outros trabalhos relacionados com o NTP que sugerem modificações e aperfeiçoamentos no protocolo.

O NTP foi proposto para operar em sistemas de propósito geral e com os mais variados sistemas operacionais. Sum-Mi Jun et al. descrevem, em seu estudo sobre um método de sincronização do tempo para o NTP [10], as restrições que esses sistemas operacionais impõem para a rede de sincronização. O trabalho sugere diversos aperfeiçoamentos para o *kernel* dos sistemas operacionais para prover características de tempo real, e um modelo de gerenciamento de tempo adequado, visto que as características do protocolo NTP implicam em várias chamadas de sistema, que devem ter alta prioridade. O trabalho apresenta uma arquitetura para implementação do NTP em sistemas embutidos com alta precisão.

David L. Mills et al. também descrevem problemas na arquitetura dos sistemas operacionais em um trabalho que apresenta um software para aumentar a resolução do protocolo NTP em uma grande variedade de sistemas. O software é chamado de *nanokernel* e possibilita maior exatidão relativa à origem do tempo, com precisão de um ou dois microsegundos, o que representa uma melhora significativa se comparado com sistemas operacionais comuns que mantêm essa precisão em centenas de microsegundos [24].

Esses trabalhos mostram que diversas melhorias podem ser aplicadas atualmente na rede de sincronização do NTP para melhorar a qualidade da sincronização realizada. No trabalho de Masato Tsuru et al. [41] é discutida a fragilidade do NTP ao estimar o *offset* correto em caminhos assimétricos. O problema de caminhos assimétricos aparece quando a largura de banda é diferente em cada sentido ou quando a rota escolhida em cada sentido apresenta diferentes distâncias. A assimetria de caminhos na Internet é cada vez mais comum devido à popularização de acessos com tecnologia xDSL e *cable modem*, dentre outras. O trabalho mostra uma técnica para estimar o *offset* do relógio e da frequência corretamente, em caminhos com largura de banda assimétricos, e cita a necessidade de pesquisa para o problema de assimetria com atrasos, o que geralmente ocorre em caminhos que dependem de satélites.

2.5 Redes Grandes e Complexas

Os mais diversos sistemas podem ser representados por grafos, onde os nodos representam o conjunto de elementos desse sistema e as arestas alguma relação ou interação entre os elementos. Um grafo pode ser utilizado para uma representação simples, como uma pequena rede de computadores em uma empresa ou para representar redes com milhares de nodos e arestas, como é o caso da Internet ou outros tipos de redes não computacionais, como a sociedade em que vivemos ou o sistema nervoso humano. Redes com milhares de nodos e arestas são conhecidas como redes “complexas”.

Um estudo realizado em 1959 [6] modelou as redes complexas como *redes aleatórias*. Uma rede aleatória é qualquer rede que não pode ser representada ou descrita em termos simples. Seguindo esse modelo, um nodo de uma rede complexa pode ter uma aresta com qualquer outro nodo com a mesma probabilidade. Dessa forma, a maioria dos nodos desse modelo tem aproximadamente o mesmo número de arestas e raramente existe um nodo com número de arestas muito acima da média. Devido ao tamanho e a complexidade da interação dos nodos dessas redes, a topologia desses sistemas é pouco conhecida.

Um novo estudo de 1999 [1], que observa várias redes complexas, contesta o modelo seguido até então e sugere um novo modelo chamado *livre de escala*. Nesse modelo, é observado que redes reais possuem nodos com poucas arestas, co-existindo com nodos com uma grande quantidade de arestas, chamados de *hubs*. Diferentemente do modelo anterior, onde os nodos são conectados de forma aleatória e uniforme, a maioria dos sistemas reais demonstra que alguns nodos possuem uma preferência para conexão. A falta de um pico na distribuição dos elementos do modelo proposto, sugere que a probabilidade $P(k)$ de interação com k nodos da rede obedeça uma *power-law*, tal como $P(k) \simeq k^{-\gamma}$.

O modelo livre de escala é dirigido por dois mecanismos: *crescimento* da rede e *conexão preferencial*. Esses mecanismos, combinados, explicam a existência de nodos que atuam como *hubs*. Uma rede inicia com poucos nodos e cresce com a adição de novos nodos. Nodos mais novos ganham conexões com menos probabilidade. Se um nodo é o primeiro na rede, todos os outros nodos têm maior probabilidade de conectar-se a ele. Dessa forma, a lei de crescimento oferece uma vantagem para os nodos mais antigos.

Devido à conexão preferencial, um nodo novo na rede prefere conectar-se ao nodo com maior número de conexões. Dessa forma, a probabilidade de um nodo ter uma nova conexão é proporcional à quantidade de conexões que o mesmo possui. Esse fenômeno foi observado em redes reais e é conhecido como *rich-get-richer*. Este conceitos são também aplicados para a rede de sincronização da Internet, como será apresentado no capítulo 5.

Capítulo 3

Coleta e Organização dos Dados

O capítulo anterior apresentou uma descrição do protocolo NTP. Vimos que o protocolo foi desenvolvido para operar numa rede heterogênea, adaptar-se facilmente a falhas e fazer automaticamente a seleção da melhor fonte de tempo. O uso do protocolo vem aumentando com o crescimento da Internet e a necessidade cada vez maior de manter os sistemas sincronizados. Alguns estudos foram realizados anteriormente entre 1989 e 1999 na rede de sincronização NTP. No entanto, nos últimos seis anos, a Internet cresceu e evoluiu em vários aspectos, desde o número de usuários e máquinas, até a infra-estrutura de rede. Assim, consideramos oportuna a realização de um novo estudo em 2005.

Para realizar o estudo proposto foi utilizado um *spider*. Genericamente, um *spider*, também conhecido como robô, é um programa que realiza consultas a um grande número de *hosts* na Internet para coletar informações. O *spider* NTP percorreu o grafo da rede de sincronização do NTP, pesquisou e armazenou em uma estrutura de dados as informações específicas do protocolo que estão disponíveis para cada *host*.

Este capítulo apresenta a coleta e a organização dos dados. As seções deste capítulo estão organizadas como descrito a seguir. A Seção 3.1 descreve os detalhes do *spider* NTP implementado para a coleta dos dados; a Seção 3.2 descreve em detalhes as consultas realizadas pelo *spider* aos *hosts* da rede NTP; a Seção 3.3 discute a necessidade de uma segunda coleta de dados e como foi esse procedimento; a Seção 3.4 detalha como os dados coletados foram organizados para a análise; por fim, a Seção 3.5 discute as questões de

segurança relacionadas a uma coleta dessa amplitude na Internet.

3.1 O *Spider* NTP

O *spider* NTP foi implementado em linguagem Java e é similar ao utilizado no estudo realizado em 1999 [26]. A nova implementação permite suporte para as duas famílias do protocolo IP, a tradicional versão 4 e a versão 6. Além disso, foi adicionado um filtro para evitar consultas em redes pré-determinadas como, por exemplo, redes privadas ou redes nas quais o administrador não desejasse receber qualquer consulta em seus servidores.

Para permitir agilidade na coleta de dados, o *spider* permite a execução de várias *threads*, ou seja, trabalha com diversas instâncias coletando e armazenando dados em paralelo. A quantidade de instâncias é determinada por um parâmetro informado na linha de comando do programa. Outros parâmetros do programa na linha de comando são dois nomes de arquivos. O primeiro arquivo contém a lista inicial de servidores a serem explorados. No segundo arquivo serão gravados os endereços dos servidores já explorados. A diferença entre o conjunto de endereços do primeiro arquivo e do segundo arquivo determina a próxima consulta ou se o trabalho completo já foi realizado. O arquivo com os dados a serem explorados é incrementado com dois valores quando um novo endereço é descoberto. O primeiro valor é o endereço IPv4 ou IPv6 do servidor e o segundo valor é o *stratum*, que é utilizado para priorizar as consultas dos servidores mais próximos da raiz na árvore de sincronização.

3.2 A Primeira Coleta de Dados

A coleta de dados teve como ponto de partida um conjunto inicial de 263 servidores públicos *stratum* 1 e 2, cujos endereços estão disponíveis em [28]. O *spider* foi executado a partir de uma única máquina, um computador Pentium 4 CPU 2.80GHz, com 512MB de memória RAM, sistema operacional Debian GNU/Linux, conectado à rede gigabit do Ponto de Presença da Rede Nacional de Ensino e Pesquisa no Paraná, PoP-PR/RNP. O PoP-PR está conectado ao backbone da RNP e a outros backbones comerciais com

enlaces de alta velocidade.

Para efetuar a coleta, o *spider* utilizou o programa `ntpd` que faz parte do conjunto de programas da implementação oficial do NTP. A RFC 1305, que especifica o Network Time Protocol versão 3, descreve um tipo de mensagem do NTP que utiliza pacotes modo 7. Esses pacotes, descritos como reservados para uso privado, são utilizados pelo programa `ntpd`. Apesar do uso privado desse tipo de mensagem, muitos servidores podem responder às consultas feitas. Por meio de mensagens do `ntpd` é possível obter a lista de monitoramento de cada servidor, com os *hosts* que fizeram alguma consulta recentemente. Essa lista é fundamental para a descoberta de novos servidores.

A primeira coleta de dados ocorreu no período de 30/08 a 05/09/2005. Nesta etapa, o *spider* NTP foi lançado para trabalhar com aproximadamente 80 *threads*. O seguinte comando foi executado para cada endereço IP:

```
./ntpd -n -c sysinfo -c peers -c monlist <endereço IP>
```

O comando `sysinfo` retorna os valores das variáveis do próprio sistema. A opção `peers` retorna a lista de associações mantidas pelo *host* e a opção `monlist` retorna a lista de monitoramento com informações dos *hosts* que fizeram alguma consulta recentemente no servidor. Foram obtidos desses dados 1.278.834 endereços IPs únicos, já excluindo os endereços eliminados pelo filtro do *spider* NTP. Destes, apenas 148.157 *hosts* responderam de alguma forma à consulta feita.

3.3 A Segunda Coleta de Dados

Uma verificação nos dados coletados mostrou que os dados de pelo menos 20.000 *hosts* estavam aparentemente corrompidos. Para verificar essa evidência foi realizado um teste sobre uma pequena amostra desses *hosts*, cujo resultado indicou que esses *hosts* não respondiam apropriadamente a solicitações feitas com pacotes modo 7 pelo programa `ntpd`. Esse teste foi feito com o programa `ntp`, que também faz parte do conjunto de programas da implementação oficial do protocolo. O `ntp` utiliza pacotes modo 6, descritos na RFC 1305 como sendo de mensagens de controle. Tal comportamento nos encorajou a fazer uma nova coleta nos endereços identificados utilizando esse programa.

Para tanto, ao invés de adaptar o *spider* NTP implementado em Java para trabalhar com o programa `ntpq`, foi desenvolvido um script para fazer a consulta utilizando esse programa nos *hosts* explorados anteriormente. Desta forma, uma segunda consulta foi realizada no período de 26 a 30/09/2005. Nesta etapa todos os 1.278.834 *hosts* identificados anteriormente foram novamente consultados, para verificação dos dados, utilizando o seguinte comando de consulta:

```
./ntpq -n -c associations -c ‘‘readvar 0’’ -c allreadvar
```

O comando `readvar 0` retorna os valores das variáveis do próprio *host*, enquanto a opção `allreadvar` retorna os dados de todas as associações de um *host*. Esta opção foi implementada no `ntpq`. Nesta etapa foram descobertos 11.985 novos endereços a partir das informações dos pares associados, totalizando 1.290.819 endereços IPs únicos, dos quais apenas 148.307 responderam às consultas. Após uma avaliação dos dados coletados, foram excluídos 1056 *hosts* que apresentaram problemas nos resultados das consultas, tais como registros mal formados ou truncados. Restaram, portanto, 147.251 *hosts* com informações úteis para a análise, ou seja, que responderam apropriadamente às consultas.

3.4 Organização dos Dados

Uma vez que os *hosts* encontrados na primeira coleta foram novamente consultados na segunda coleta e retornaram resultados corretos para as consultas realizadas, esses dados foram escolhidos para montar a base de dados para a análise da rede de sincronização do NTP. Os dados coletados somam 870Mbytes e estão armazenados de maneira simples. Para cada *host* explorado que retornou algum dado foi criado um arquivo cujo nome é o endereço IP do *host*.

Os dados coletados foram agrupados em dois grandes grupos. O primeiro grupo contém informações de todas as variáveis dos sistemas consultados, ou seja, informações locais de cada servidor que foram obtidas com a opção `readvar 0` do comando `ntpq`. O segundo grupo contém informações das variáveis dos pares associados a esses servidores, que foram obtidas com a opção `allreadvar` do comando `ntpq`. Essas informações foram importadas para duas tabelas em um banco de dados MySQL, denominadas *server* e *peer*,

Campo	Tipo	Descrição
server	char	Endereço IP do <i>host</i>
status	int	Variável <i>status</i> do <i>host</i>
proc	char	Identificador do modelo do processador do sistema
system	char	Identificador do sistema operacional do <i>host</i>
leap	int	Identificador para inserção/remoção do <i>leap second</i>
<i>stratum</i>	int	Valor do <i>stratum</i> do sistema
rootdelay	float	Atraso para a fonte de tempo primária (raiz)
rootdispersion	float	Erro relativo à fonte de tempo primária (raiz)
peer	int	Identificador para a associação designada como par de sincronização
refid	char	Código de identificação do relógio de referência local
reftime	char	A estampa do tempo no instante da última atualização do relógio local
clock	char	A estampa do tempo do relógio local
poll	int	Indicador do intervalo mínimo de troca de mensagens
state	int	Estado da disciplina do relógio local
offset	float	Valor do <i>offset</i> do sistema
frequency	float	Valor do <i>offset</i> de frequência do sistema
jitter	float	Valor do <i>jitter</i> do sistema
stability	float	Estabilidade do relógio local

Tabela 3.1: Campos da tabela *server*.

respectivamente.

A tabela *server* foi indexada pelo endereço IP do servidor, enquanto a tabela *peer* foi indexada pelo endereço IP do servidor e pelo identificador da associação que o servidor mantinha com o *peer*, permitindo correlacionar as tabelas para buscar as informações. Os campos de cada tabela, os tipos de dados utilizados e suas descrições são mostrados na Tabelas 3.1 e 3.2.

3.5 Questões de Segurança

Um dos problemas principais deste tipo de coleta de dados é a segurança. As consultas podem ser confundidas com acessos indevidos, embora não sejam, pois o protocolo NTP permite este tipo de consulta. Para minimizar este problema, antes de se iniciar a coleta dos dados foi construída uma página Web¹ com versões em inglês e português, contendo diversas informações sobre o estudo realizado, como os endereços de origem das consultas, o período de realização do estudo, o tipo de consulta executado, o endereço de contato dos autores, dentre outras informações. Uma entrada reversa no DNS sobre o domínio utilizado (ntpsurvey.arauc.br) foi adicionada, para permitir que os administradores de

¹<http://www.ntpsurvey.arauc.br>

Campo	Tipo	Descrição
server	char	Endereço IP do sistema
assid	int	Identificador interno para esta associação
status	int	Variável <i>status</i> do par
srcaddr	char	Endereço de origem das mensagens NTP
srcport	int	Porta de origem das mensagens NTP
dstaddr	char	Endereço de destino das mensagens NTP
dstport	int	Porta de destino das mensagens NTP
leap	int	Identificador para inserção/remoção do <i>leap second</i>
stratum	int	Valor do <i>stratum</i> do par
precision	float	Atraso do par para a fonte de tempo primária
rootdispersion	float	Erro relativo do par para a fonte de tempo primária na raiz
refid	char	Código de identificação do relógio de referência local do par
reach	int	Registrador de deslocamento que determinada a alcançabilidade do par
unreach	int	Contador de intervalos que se passou desde o último pacote recebido
hmode	int	Modo de operação do servidor para esse par
pmode	int	Modo de operação do par
hpoll	int	Indicador do intervalo entre as trocas de mensagens do servidor para o par
ppoll	int	Indicador do intervalo entre as trocas de mensagens do par
flash	int	Código (para fins de depuração) do último pacote recebido
keyid	int	Identificador da chave criptográfica usada para gerar o código de autenticação
offset	float	Valor do <i>offset</i> para o par
delay	float	Valor do atraso para o par
dispersion	float	Erro relativo para o par
jitter	float	Valor do <i>jitter</i> para o par
reftime	char	A estampa do tempo no instante que o relógio do par foi atualizado pela última vez
org	char	A estampa do tempo no par quando a última mensagem foi transmitida.
rec	char	A estampa do tempo no sistema quando a última mensagem foi recebida
xmt	char	A estampa do tempo quando a mensagem foi enviada pelo sistema

Tabela 3.2: Campos da tabela *peer*.

rede localizassem a página Web quando notassem o acesso a algum nodo de sua rede, obtendo informações sobre tal acesso.

Durante a coleta dos dados foram reportados cerca de dez incidentes de segurança. A maioria das mensagens foi enviada para o endereço de contato dos prefixos IP listados no sistema *whois directory service*. Apenas duas comunicações foram feitas diretamente para o endereço de contato listado na página anunciada. Um único administrador insistiu em não querer mais receber qualquer acesso em sua rede, e então os endereços de sua rede foram adicionados ao filtro do *spider* NTP. Aparentemente, a maioria dos incidentes reportados foram gerados por ferramentas automáticas, como o IDS (*Intrusion Detection and Prevention System*). Os detalhes da coleta de dados estão descritos em [40].

Capítulo 4

Caracterização da Qualidade da Sincronização

Este capítulo descreve os resultados da caracterização da qualidade da rede de sincronização do NTP. Diversas variáveis do protocolo NTP registram a qualidade de sincronização do sistema. Neste capítulo serão discutidas as métricas *offset*, RTT, dispersão, *jitter* e estabilidade dos relógios. A caracterização dessas métricas foi feita com base em índices de tendência central e de dispersão. Para avaliar a dispersão de um conjunto de dados em torno da média foi utilizado o coeficiente de variação (COV), definido como a razão entre o desvio padrão e a média.

Para cada conjunto de dados relativo a uma métrica foram plotados a frequência de ocorrência dos valores (PDF), a frequência cumulativa (CDF) e o complemento da frequência cumulativa (CCDF). Os gráficos de frequência permitem verificar a ocorrência de um evento em uma faixa de valores. Os gráficos de frequência cumulativa definem a probabilidade de uma observação ser menor que um valor e, por fim, os gráficos de complemento da frequência cumulativa são úteis para avaliar a ocorrência de valores extremos na distribuição dos dados.

Vários resultados são comparados com dados publicados em estudos anteriores. Do nosso conhecimento, este é o quinto estudo sobre a rede de sincronização do NTP no período entre 1989 e 2005. As publicações anteriores e suas respectivas datas estão citadas

Ano	Referência Bibliográfica	Tamanho da Rede	Respostas
1989	[20]	8.455	946
1993	[9]	15.000	7.251
1995	[25]	—	38.722
1999	[26]	647.401	175.527
2005	Este trabalho	1.290.819	147.251

Tabela 4.1: Crescimento da rede NTP ao longo do tempo.

Código	Descrição da Fonte de Tempo
0	Não especificada
1	Calibrada atomicamente
2	Rádio VLF ou LF
3	Rádio HF
4	Rádio UHF
5	Local
6	NTP
7	UDP/TIME
8	Relógio de pulso
9	Modem/telefone
32	Sinal PPS

Tabela 4.2: Códigos de origem de sincronização da variável de sistema *status*

na Tabela 4.1. A coluna “Ano” desta Tabela refere-se ao ano da coleta de dados, não ao ano da publicação. Várias tabelas apresentadas neste capítulo utilizam o ano da coleta de dados como referência para a respectiva publicação.

Nas consultas realizadas foi utilizada a variável *status* do sistema e dos pares. Embora o nome da variável seja o mesmo nos dois conjuntos de dados, o significado é diferente. A variável *status* do sistema mantém, dentre outras informações, a origem de sincronização, conforme a codificação descrita na Tabela 4.2. Esta codificação será utilizada frequentemente nessa análise para selecionar os servidores que possuem uma fonte de tempo especificada. A variável *status* dos pares mantém, dentre outras informações, o status do par, determinado pelos algoritmos do NTP, conforme a codificação mostrada na Tabela 4.3. Esta codificação será usada frequentemente na análise para encontrar o par de sincronização escolhido pelo sistema.

As demais seções deste capítulo estão organizadas como descrito a seguir. A Seção 4.1 descreve o tamanho da rede NTP alcançada; a Seção 4.2 mostra como está constituída

Código	Descrição
0	Par rejeitado por ser inalcançável ou por ter sido identificado um erro grosseiro
1	Par descartado pelo algoritmo <i>clock selection</i> como sendo um <i>falseticker</i>
2	Par aprovado pelo algoritmo <i>clock selection</i> porém apresenta uma alta distância de sincronização
3	Par descartado pelo algoritmo <i>clustering</i> por apresentar uma alta dispersão ou alto <i>jitter</i>
4	Par sobrevivente e candidato ao algoritmo <i>combining</i>
5	Par sobrevivente mas não apresenta uma boa distância de sincronização
6	Par utilizado para sincronização
7	Par derivado de uma fonte de sincronização de pulso – PPS

Tabela 4.3: Códigos para a variável *status* dos pares

a hierarquia de servidores NTP; a Seção 4.3 discute os resultados da caracterização da qualidade de sincronização obtida pelo NTP; a Seção 4.4 apresenta a correlação entre algumas variáveis do NTP. Por fim, a Seção 4.5 faz a síntese do capítulo.

4.1 Tamanho da Rede de Sincronização

A Tabela 4.1 apresenta dados sobre o crescimento da rede de sincronização, publicados em estudos anteriores, bem como os resultados obtidos pela caracterização apresentada neste trabalho. A coluna “Tamanho da Rede” indica o número de *hosts* que foram descobertos, enquanto a coluna “Respostas” indica o número de *hosts* que efetivamente responderam às consultas do protocolo NTP.

O tamanho da rede NTP cresceu cerca de três ordens de grandeza desde o primeiro estudo. Comparando com o último estudo, realizado em 1999, o crescimento foi de quase 100%. No entanto, neste mesmo período, a fração de *hosts* que respondeu às consultas diminuiu de 27% para 11% do total de *hosts* descobertos. Ainda assim, consideramos que a quantidade de dados para essa análise é suficiente para o estudo.

Cerca de 89% dos *hosts* descobertos não puderam ser alcançados ou não responderam às consultas por várias razões. Por exemplo, alguns *hosts* podem ter seu acesso bloqueado por um *firewall*. Com a popularização da rede e o aumento de incidentes de segurança, muitos administradores têm optado por evitar manter disponível algum serviço público, pois a possibilidade de acesso pode aumentar a vulnerabilidade do sistema. Além disso, cada endereço IP foi consultado apenas uma vez, e as consultas ou as respostas podem ter sido perdidas devido a falhas ou a congestionamento na rede, uma vez que elas são implementadas com o protocolo UDP. Outro motivo pode ser atribuído ao fato de que

muitos dos *hosts* listados podem ter sido adicionados pelo *spider* por terem feito uma consulta a um servidor NTP, mas esses *hosts* não mantêm um *daemon* NTP executando continuamente. Uma outra explicação sugere que alguns servidores podem não ser capazes de responder às consultas executadas, possivelmente devido a diferenças entre as versões do protocolo e do programa de consulta.

4.2 Caracterização da Hierarquia da Rede de Sincronização

A principal característica da topologia da rede NTP é a hierarquia dos servidores, definida pelo *stratum*. A Tabela 4.4 mostra a frequência de servidores por *stratum*. A consulta mostrada abaixo foi realizada no banco de dados para obter a quantidade de servidores que estão consultando o tempo em alguma fonte conhecida pelo protocolo NTP, para cada valor de *stratum*:

```
select stratum, count(1) from server where ((status >> 8) & 63)1 != 0 group by stratum;
```

A maior parte dos servidores atua como *stratum* 2, 3 e 4, sendo os servidores *stratum* 3 os mais comuns. Em números absolutos, foram identificados 533 servidores *stratum* 1, 28.687 servidores *stratum* 2 e 82.145 servidores *stratum* 3. Este resultado é explicado pela forma como estão distribuídos os servidores NTP na Internet. Servidores *stratum* 1, que sincronizam seus relógios diretamente com uma fonte de tempo precisa e confiável, são fonte de sincronização para servidores *stratum* 2, que estão localizados normalmente no núcleo da rede. Estes servidores, por sua vez, são consultados por diversos servidores *stratum* 3 localizados nas bordas, que distribuem a informação de tempo para as redes locais. Resultado similar foi observado em [26].

Em geral, o que se observa na Internet é que os servidores *stratum* 1 públicos estão disponíveis apenas para consultas de *hosts* que atuarão com *stratum* 2 [28]. A lista desses servidores disponíveis é pequena e o controle de acesso é mais restrito que para os servidores *stratum* 2. Quando um administrador de sistema busca fontes de sincronização na Internet, acaba optando por sincronizar sua máquina com servidores *stratum* 2, que

¹Deslocamento de 8 posições a direita e operação que retorna os bits do código da variável *status*

<i>stratum</i>	1	2	3	4	5	6	7 – 10	11	12–16
Quantidade	533	28687	82145	24698	1786	318	298	569	253
Fração (%)	0,38	20,60	58,98	17,73	1,28	0,23	0,21	0,41	0,18

Tabela 4.4: Distribuição dos servidores por *stratum*.

estão próximos geograficamente. Assim, os servidores *stratum* 2 são responsáveis pela sincronização de milhares de *hosts* que atuam como *stratum* 3.

Nota-se uma quantidade considerável de servidores *stratum* 11 na Tabela 4.4. Uma análise mais detalhada indica que muitos administradores configuram o servidor local como relógio de referência, atribuindo valor 10 ao *stratum*. Dos 569 servidores *stratum* 11, 505 possuem o relógio de referência configurado com o endereço de *loopback*. Essa configuração permite que outros *hosts* utilizem este servidor como fonte de sincronização quando servidores de *stratum* menor não estão disponíveis.

A Tabela 4.5 apresenta a distribuição dos servidores por *stratum* ao longo dos anos. Em todas as caracterizações realizadas, a porcentagem de servidores *stratum* 1 é pequena mas, ainda assim, pode-se observar uma sensível diminuição nesta porcentagem nos últimos anos. Esta tendência pode ter duas explicações. A primeira explicação refere-se ao custo de aquisição e à dificuldade de manutenção de um servidor *stratum* 1. A segunda explicação é devida a um modelo de crescimento e evolução da Internet denominado “livre de escala” [1, 2], e suas derivações [5, 31]. Um dos princípios deste modelo indica que os nodos se conectam preferencialmente a nodos já bastante conectados, um fenômeno denominado *rich-get-richer*. Este modelo relaciona o tempo de vida de um nodo da rede ao número de conexões para o nodo. Nodos altamente conectados (*hubs*) não morrem e raramente perdem suas conexões. Segundo este modelo, o surgimento de novos servidores *stratum* 1 seria desencorajado, pois, pelo menos a princípio, eles teriam poucos clientes. É importante ressaltar que o pequeno número de servidores *stratum* 1 reforça a grande importância do comportamento desses servidores para a rede de sincronização do NTP.

Observa-se também, ainda na Tabela 4.5, que a quantidade de servidores *stratum* 3, que era próxima dos 47% nos estudos anteriores, aumentou para quase 59%. Este dado pode ser uma evidência de que, frente à necessidade crescente de sincronização com

<i>stratum</i>	1993	1995	1999	2005
1	0,91	1,58	0,55	0,38
2	20,36	31,84	15,29	20,60
3	46,53	47,28	48,61	58,98
4	27,60	16,17	21,84	17,73
5	0,52	2,27	4,06	1,28
6	0,08	0,43	0,94	0,23
7 a 15	0,50	0,43	0,55	0,80
16	3,50	0	2,77	0

Tabela 4.5: Evolução da distribuição (%) de servidores por *stratum* ao longo do tempo.

exatidão, mais pessoas estão procurando sincronizar suas máquinas com servidores NTP mais próximos da raiz.

4.3 Caracterização da Qualidade da Sincronização

Nesta seção são discutidas métricas que caracterizam a qualidade da sincronização obtida e mantida pela rede NTP. O NTP calcula as métricas *clock offset*, RTT, *jitter* e dispersão. Todas estas medidas são relativas a um relógio de referência. Os resultados destas medidas são discutidos nas subseções a seguir.

4.3.1 Análise do Offset

O *clock offset*, ou simplesmente *offset*, é o valor do ajuste necessário ao relógio local para colocá-lo em fase com o relógio de referência, ou seja, é a diferença entre esses relógios. O *clock offset* é o indicador mais importante da qualidade da sincronização. Quanto menor o *offset*, melhor é a sincronização.

Duas medidas de *offset* foram extraídas dos dados, com as consultas mostradas abaixo:

```
select abs(offset) from server where ((status >> 8) & 63) != 0;
```

```
select abs(offset) from peer where ((status >> 8) & 7) = 6;
```

A primeira medida é o *offset* da variável do sistema, fornecida por cada servidor, e representa o *offset* utilizado para fazer o ajuste no relógio local. A segunda informação de *offset* é obtida nas variáveis dos pares, da associação que está servindo como par de

Métrica	<i>Offset</i> (ms)	
	Variável de Sistema	Variável do Par
Menor	0	0
Maior	1,49e+08	1,12e+12
Média	4.217	8,23+06
Mediana	0,7	0,8
COV	147	370
Primeiro Quartil	0,18	0,22
Terceiro Quartil	2,91	3,59
97º Percentil	86,67	124,54

Tabela 4.6: Estatísticas para a variável *offset* do sistema e do par

sincronização, e representa o valor de *offset* coletado mais recentemente relacionado com um baixo atraso (RTT).

A Tabela 4.6 sumariza os valores absolutos das duas medidas. O menor valor encontrado nos dois conjuntos de dados é zero. Isto acontece quando o relógio local está totalmente em fase com o relógio de referência. O maior valor encontrado para os *offsets* está na ordem de centenas de milhares de milissegundos, indicando que nesses sistemas o relógio não está sincronizado com sua fonte de referência. As médias estão elevadas devido à grande variabilidade nos dados. As medianas indicam que pelo menos a metade dos sistemas está com menos de 1 milissegundo de diferença para a hora certa. O primeiro quartil mostra que muitos sistemas estão com valores bem menores. O terceiro quartil indica a que a maioria dos sistemas sincroniza com exatidão na ordem de poucos milissegundos. O 97º percentil mostra que apenas uma pequena fração dos sistemas apresenta *offset* muito distante do seu par de sincronização.

O gráfico de probabilidade do *offset* dos servidores e dos pares é apresentado à esquerda da Figura 4.1. Nota-se que a frequência do *offset* dos servidores e dos pares é bastante parecida. Observa-se poucos sistemas com *offset* até 0,01 ms e também poucos sistemas com valores de *offset* acima de 100 ms. A maior concentração é de sistemas com *offset* entre 0,1 ms a 10 ms.

As curvas da frequência cumulativa dos *offsets* dos servidores e dos pares são apresentadas no centro da Figura 4.1. As curvas são bastante semelhantes e é possível notar que 95% dos dados possui *offset* de no máximo 50 ms.

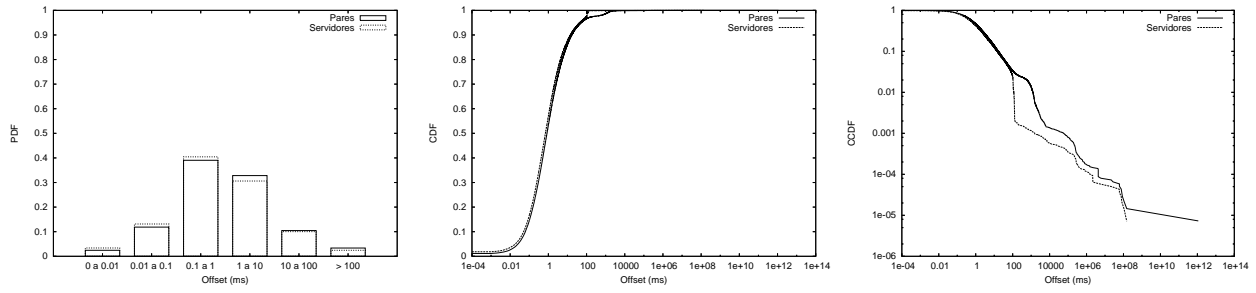


Figura 4.1: PDF, CDF e CCDF do *offset* em relação aos pares e aos servidores.

As curvas do complemento da frequência cumulativa, apresentadas no gráfico à direita da Figura 4.1, estão sobrepostas até o valor de *offset* de aproximadamente 100 ms, ponto em que se observa uma queda abrupta na curva de *offset* dos servidores. A partir deste ponto as duas curvas se distinguem, sendo o *offset* do conjunto de servidores menor do que o *offset* dos pares. Isto é explicado pelo fato que o NTP define um limite superior de 128 ms para o ajuste do relógio. Uma vez que o relógio do *host* foi sincronizado com o relógio de referência, dificilmente o *offset* ultrapassará este limiar, mesmo em condições de congestionamento na rede e outros atrasos. Uma amostra que excede esse valor é descartada, isto é, não é considerada pelos algoritmos de ajuste de relógio. Porém, se o *offset* permanecer acima do limiar por um longo tempo, em geral 900 s, o relógio local é atualizado para o tempo indicado pelo relógio de referência.

Os valores de *offset* apresentam grande variabilidade. A cauda longa à direita, observada em ambas as curvas do gráfico CCDF, indica que uma pequena fração dos *hosts* apresenta uma grande diferença em relação aos relógios de seus pares e servidores, podendo esta diferença chegar a horas e mesmo dias. No entanto, para a variável do sistema, apenas 0,2% dos dados apresenta valores de *offset* maiores que 128 ms.

A Tabela 4.7 apresenta a evolução dos valores de *offset* nos últimos dez anos. Esta tabela mostra a fração de servidores com *offset* maior que 128 ms, bem como a mediana e a média dos valores de *offset* menores que 128 ms. A comparação com estudos anteriores revela melhoria considerável na qualidade da manutenção de tempo. Uma fração cada vez maior de servidores faz ajustes cada vez menores em seus relógios. Este valores demonstram que a sincronização mantida pela rede NTP atual é significativamente melhor do que no passado.

Ano	<i>offset</i> > 128 ms	Mediana	Média
1995	9,9%	20,1	28,7
1999	3,0%	1,8	8,2
2005	0,2%	0,7	7,0

Tabela 4.7: Evolução da métrica *offset* ao longo do tempo.

As consultas retornam o valor absoluto do *offset* para análise. No entanto, o *offset* pode ser negativo, o que indica que o relógio local teria que fazer um ajuste para um tempo no passado, o que violaria a condição de monotonicidade. Quase 50% dos valores de *offset* são negativos, o que significa que cerca de metade dos ajustes de relógios seria para um tempo no passado. Porém, o NTP busca fazer o ajuste no relógio em passos pequenos a cada segundo, modificando a frequência do oscilador. Apenas se o valor de ajuste for elevado, em geral maior que 128 ms, o NTP faz a correção em uma única etapa, mesmo sendo um tempo no passado e, assim, violando o princípio da monotonicidade.

4.3.2 Atraso para o Par de Sincronização e a Raiz

A métrica de atraso ou distância entre dois nodos da rede é o RTT. O protocolo NTP calcula o atraso para cada associação e a distância para a raiz, que representa o valor acumulado do RTT para a origem da referência de tempo primária na sub-rede de sincronização.

O RTT é uma medida importante para a sincronização via NTP, uma vez que a exatidão da estampa de tempo de um servidor depende de um RTT pequeno para seu par de sincronização e para a raiz. Em particular, o NTP assume que o caminho de rede entre *hosts* é simétrico. Esta premissa pode inserir um erro que será tanto menor quanto menor for o RTT.

As consultas mostradas abaixo foram efetuadas no banco de dados para se obter os valores de atraso para o par de sincronização e para a raiz na árvore de sincronização através da variável do par *delay* e da variável do sistema *rootdelay*, respectivamente:

```
select abs(delay) from peer where ((status >> 8) & 7) = 6;

select abs(rootdelay) from server where ((status >> 8) & 63) != 0;
```

Métrica	RTT (ms)	
	Par	Raiz
Menor	0	0
Maior	5.000	11.885
Média	28,99	64,24
Mediana	4,19	29,59
COV	2,71	1,81
Primeiro Quartil	0,9	10,62
Terceiro Quartil	23,21	73,05
97º Percentil	196,13	313,34

Tabela 4.8: Estatística do RTT para o par de sincronização e para a raiz.

Um valor negativo de atraso para o par pode ser calculado pelo NTP, dependendo da diferença entre os relógios do par de sincronização e do cliente. A consulta do atraso para o par é efetuada com valor absoluto, permitindo assim a comparação entre as duas métricas.

A Tabela 4.8 sumariza os valores das medidas de atraso para o par e para a raiz. O menor valor encontrado nos dois conjuntos de dados é zero. Isto acontece quando o relógio local está diretamente conectado com o relógio de referência. O maior valor encontrado para o *atraso* é da ordem de alguns segundos, indicando que alguns sistemas estão utilizando um caminho longo ou congestionado até o relógio de referência. Em geral, o atraso para a raiz é maior do que para o par. Isto acontece devido à hierarquia do NTP e à pouca quantidade de servidores primários. Assim, os caminhos até a fonte primária de tempo são longos. Apesar disso, o RTT médio para a raiz é apenas 64 ms e a mediana é aproximadamente 29 ms. Para o par, esses valores são ainda mais baixos, a média é 29 ms e a mediana 4 ms. O primeiro quartil indica que alguns sistemas enfrentam caminhos curtos até o relógio de referência. O terceiro quartil indica caminhos mais longos, da ordem de dezenas de ms, e o 97º percentil indica que poucos sistemas tem seu caminho para o relógio de referência longo ou congestionado (centenas de ms). O COV é baixo nos dois conjuntos de dados, indicando pouca variabilidade nos atrasos.

O gráfico de frequência dos atrasos para o par e para a raiz é apresentado à esquerda da Figura 4.2. Nota-se que RTTs menores são mais frequentes para o par de sincronização. Isto indica que a maior parte dos *hosts* sincroniza com pares fisicamente próximos. O

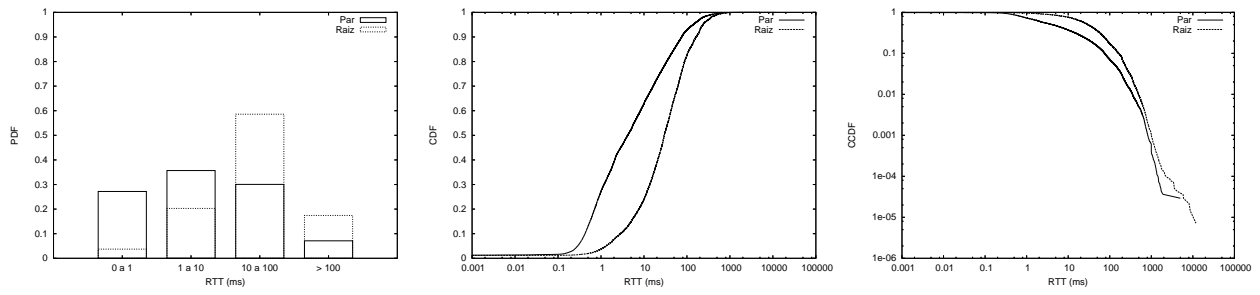


Figura 4.2: PDF, CDF e CCDF do RTT para o par de sincronização e para a raiz.

Ano	Mediana	Média	COV
1995	118	186	—
1999	32	33	3,48
2005	4,2	29	2,71

Tabela 4.9: RTT (ms) para o par de sincronização em diversos estudos.

RTT para a raiz ocorre com mais frequência nas classes com valores mais elevados. Em geral, isso ocorre devido à pouca quantidade de servidores operando como *stratum* 1. Dessa forma, é necessário um caminho mais longo para alcançá-los.

As curvas da frequência cumulativa dos atrasos para o par e para a raiz são apresentadas no centro da Figura 4.2. O RTT para o par apresenta atraso inferior ao RTT para a raiz. Para o par, 92% dos dados tem RTT inferior a 100 ms, e para a raiz, 82% é menor que 100 ms. Nas curvas do complemento da frequência cumulativa, apresentadas à direita da Figura 4.2, observa-se que cerca de 5,6% dos servidores apresentam atraso para a raiz acima de 500 ms e 0,4% tem atraso para a raiz acima de 10 segundos. Cerca de 0,8% dos servidores tem atraso para o par de sincronização acima de 500 ms. Estes números sugerem que uma pequena porcentagem dos *hosts* sincroniza com pares distantes, via caminhos longos ou congestionados.

Comparações da média e da mediana do RTT para o par de sincronização e para a raiz, em diversos estudos, são apresentadas, respectivamente, nas Tabelas 4.9 e 4.10. Os resultados indicam um melhoramento significativo na latência dos caminhos da rede NTP e, portanto, da Internet. Esta melhoria na infra-estrutura da rede pode explicar a melhoria na exatidão dos relógios dos servidores, demonstrada na seção anterior.

Ano	Mediana	Média	COV
1993	100	160	1,05
1999	47	84	1,86
2005	30	64	1,81

Tabela 4.10: Distância para a raiz (tempo em ms) em diversos estudos.

<i>Stratum</i>	1993		1999		2005	
	média	COV	média	COV	média	COV
1	105	1,06	80	2,34	55	1,74
2	42	1,76	29	3,52	24	2,79
3	36	1,72	15	4,47	18	4,83
4	42	0,45	12	3,17	17	3,47
5	50	0,38	3	5,33	9	8,89

Tabela 4.11: RTT (ms) para o par de sincronização por *stratum* em diversos estudos.

4.3.3 Atraso para o Par de Sincronização por *stratum*

A Tabela 4.11 apresenta valores de RTT para o par de sincronização, classificando os resultados de acordo com o *stratum* do par de sincronização. Observa-se que os maiores RTTs são para servidores cujos pares de sincronização são *stratum* 1, em todos os estudos realizados. Os valores de RTT para servidores nos demais níveis são significativamente menores.

Este resultado demonstra o arranjo hierárquico do NTP no ambiente real, com servidores *stratum* 1 mais distantes, seguidos por servidores de *strata* maiores, que estão mais próximos dos pares de sincronização. A diminuição do RTT médio para servidores de *strata* 1 e 2 ao longo do tempo também indica um melhoramento na infra-estrutura da rede. Os valores maiores para o COV nos últimos dois estudos podem ser uma evidência da heterogeneidade dos sistemas e redes que compõem a Internet. As redes ficaram mais rápidas, porém mais heterogêneas.

4.3.4 Dispersão para a Raiz e para o Par de Sincronização

O protocolo NTP calcula também a dispersão para raiz e para o par de sincronização. A dispersão indica o erro máximo em relação à fonte de referência (raiz ou par), e é

Métrica	Dispersão (ms)	
	Par	Raiz
Menor	0	0
Maior	16.000	5,66e+07
Média	30,64	5.629,48
Mediana	2,78	41,87
COV	12,77	72,10
Primeiro Quartil	0,32	22,66
Terceiro Quartil	14,84	74,61
97º Percentil	151,76	1169,30

Tabela 4.12: Estatística da dispersão para o par de sincronização e para a raiz.

medida em segundos. As consultas mostradas abaixo foram executadas no banco de dados para se obter esses valores para o par de sincronização e para a raiz na árvore de sincronização, armazenadas na variável do par *rootdispersion* e na variável do sistema *dispersion*, respectivamente:

```
select dispersion from peer where ((status >> 8) & 7) = 6;
```

```
select rootdispersion from server where ((status >> 8) & 63) != 0;
```

A Tabela 4.12 resume os valores das medidas de dispersão para o par e para a raiz. O menor valor encontrado nos dois conjuntos de dados é zero. Isto ocorreu em poucos casos (0,01% dos dados), e uma análise nos dados coletados indica relação com a implementação do NTP de alguns sistemas. O maior valor encontrado para a dispersão para o par é de 16 s. Este valor é um limiar do NTP, para indicar amostras perdidas. Para a raiz, esse valor é bem superior, indicando que alguns *hosts* da hierarquia até a fonte primária de tempo estão com a dispersão elevada. A média e o COV da dispersão para a raiz são valores elevados comparando-se com os valores para o par. A mediana para o par é 2,78 ms e para a raiz esse valor é 41,87 ms. Os quartis também indicam uma maior amplitude dos valores para a raiz. A dispersão para raiz é maior do que a dispersão para o par, pois é o erro acumulado no caminho entre o servidor e a raiz.

O gráfico de frequência das dispersões para o par e para a raiz é apresentado à esquerda da Figura 4.3. Nota-se que a dispersão para o par é mais freqüente nas primeiras classes, sendo predominante na classe com valores menores. A dispersão para a raiz é mais freqüente nas classes com valores mais elevados, porém observa-se uma menor quantidade

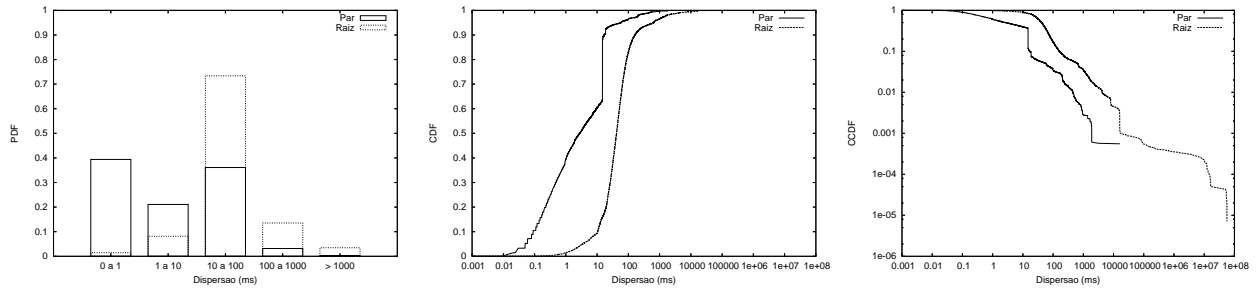


Figura 4.3: PDF, CDF e CCDF da dispersão para o par de sincronização e para a raiz.

de sistemas na última classe, que contém valores mais acentuados.

As curvas da frequência cumulativa da dispersão para o par e para a raiz são apresentadas no centro da Figura 4.3. A curva com os valores para o par apresenta dispersão inferior aos valores para a raiz. Para o par, 99% dos dados tem dispersão inferior a 1 s e, para a raiz, 96% é menor que 1 s. Nas curvas do complemento da frequência cumulativa, apresentadas à direita da Figura 4.3, observa-se que a dispersão para raiz apresenta uma cauda longa, compostas por aproximadamente 3% do conjunto de dados. A maior dispersão encontrada é de cerca de 16 horas. Para o par a dispersão máxima encontrada é o limiar de 16 s.

Uma comparação com outros estudos também revela uma melhoria nesta métrica. No estudo de 1999, a maior dispersão tinha valor superior a um ano. A cauda também era composta por cerca de 3% do conjunto de dados. Eliminando a cauda de ambos os conjuntos de dados, os valores da mediana e da média para a dispersão, nos estudos de 1999 e 2005, são, respectivamente, 39 e 88, e 40 e 74 ms. Nota-se uma diminuição nos valores da média, mas não da mediana. Este resultado pode indicar que, embora haja uma diminuição no valor médio do erro medido pela dispersão, há um limite inferior dado pelo valor da mediana. Este limite pode estar relacionado à qualidade dos osciladores dos relógios utilizados nos computadores.

4.3.5 Análise do Jitter

A versão 4 do protocolo NTP mantém uma variável chamada *jitter* para o sistema e para cada associação. Para os pares, esse valor é calculado pelo algoritmo *clock filter*

e, em seguida, utilizado pelo algoritmo *clustering* como uma métrica de qualidade que representa a variação nas amostras de *offset* para o par. Essa variação ocorre, em geral, devido a características na infra-estrutura da rede, como variação do atraso, caminhos congestionados ou assimétricos. Para o sistema, esse valor representa a melhor estimativa de erro na computação do *offset*, tendo sido determinado pelo protocolo levando em consideração não apenas a variação nas amostras do *offset* para o par de sincronização, mas também a variação do *offset* entre as associações sobreviventes ao algoritmo *clustering*. As duas medidas de *jitter* foram extraídas do banco de dados de acordo com as consultas abaixo:

```
select jitter from server where ((status >> 8 ) & 63) != 0
select jitter from peer where ((status >> 8) & 7 ) = 6
```

A Tabela 4.13 sumariza os valores da variável *jitter* do sistema e do par. O menor valor encontrado nos dois conjuntos de dados é zero. Isto acontece quando o relógio local está diretamente conectado com o relógio de referência, não sofrendo variação nas amostras de *offset* calculadas. O maior valor encontrado para o *jitter* da variável de sistema é de 842 s e, para a variável de par, 22 s. Sistemas com poucas associações, mas com *offset* muito variável, elevam o valor do *jitter*. O *jitter* médio para o par é 9,52 ms, e o *jitter* do sistema é 81 ms. Metade dos sistemas tem um *jitter* de 2,81 ms na variável de sistema, e de 0,51 ms para o par. O COV do *jitter* do sistema é de 53, e para o par é de 11. Os quartis também indicam maior amplitude dos valores para o sistema. Os valores mais elevados nos dados da variável de sistema se justificam pelo fato do protocolo NTP utilizar diversas associações para o cálculo desse *jitter*, que representa a estimativa de erro na computação do *offset*, e para o par, o *jitter* representa a variação nas amostras calculadas nas trocas de mensagens.

O gráfico de frequência do *jitter* dos servidores e dos pares é apresentado à esquerda da Figura 4.4. Nota-se que a frequência do *jitter* dos servidores é predominante nas classes com valores de 1 a 100 ms. Para os pares, a frequência é mais comum nas classes com valores entre 0 e 10 ms. Isto ocorre pois o cálculo do *jitter* do sistema considera outras associações além daquela escolhida como par de sincronização.

Métrica	<i>Jitter</i> (ms)	
	Variável de Sistema	Variável do Par
Menor	0	0
Maior	842.922	22.630
Média	80,91	9,52
Mediana	2,81	0,51
COV	53,15	11,05
Primeiro Quartil	0,7	0,12
Terceiro Quartil	11,15	2,52
97º Percentil	122,95	66,45

Tabela 4.13: Estatística do *jitter* da variável de sistema e da variável do par

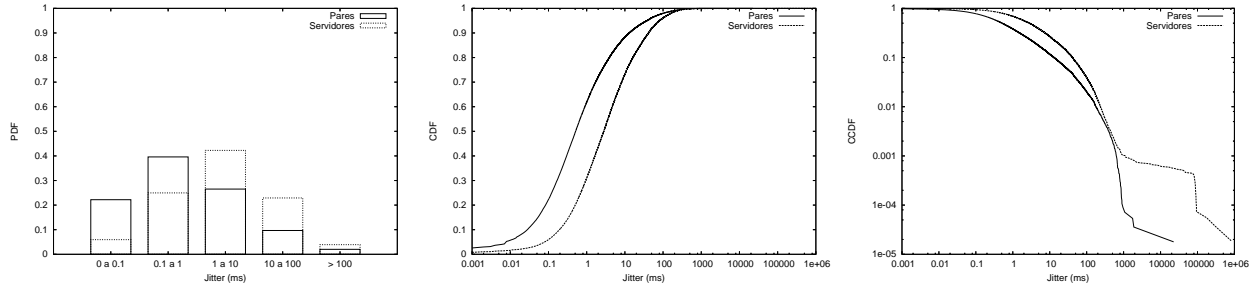


Figura 4.4: PDF, CDF e CCDF do *jitter* em relação aos pares e aos servidores

As curvas da frequência cumulativa dos valores do *jitter* dos servidores e dos pares de sincronização são apresentadas no centro da Figura 4.4. As curvas são semelhantes, porém a curva do *jitter* do sistema apresenta valores superiores ao *jitter* do par de sincronização. Isto ocorre porque o algoritmo *clock combining* utiliza amostras de associações que são estatisticamente equivalentes no cálculo do *offset* e do *jitter* do sistema. Nas curvas do complemento da frequência cumulativa, apresentadas à direita da Figura 4.4, nota-se que para uma pequena fração dos dados, 0,2%, as curvas se distanciam. Isto revela que em alguns casos o protocolo não consegue fazer a distinção entre amostras boas e ruins, indicando que, ou poucas associações são mantidas no servidor, e estas não são boas fontes de tempo, ou todas as associações mantidas estão sofrendo a influência de algum problema em comum, como, por exemplo, um caminho muito congestionado.

Como apenas as versões mais recentes do protocolo calculam a variável *jitter*, não é possível fazer uma comparação com outros estudos.

Métrica	Estabilidade do Relógio (PPM)
Menor	0
Maior	2.908.202
Média	227,75
Mediana	0,015
COV	79,85
Primeiro Quartil	0,003
Terceiro Quartil	0,239
97º Percentil	32,57

Tabela 4.14: Estatística da estabilidade dos relógios

4.3.6 Estabilidade dos Relógios

Muitos fatores influenciam a qualidade do oscilador de um computador, como temperatura, umidade, regulação da fonte, vibração, dentre outros. A variável de sistema *stability* é uma medida disponível no conjunto de variáveis locais e permite uma análise da qualidade dos osciladores. A estabilidade dos relógios é calculada como a média exponencial da diferença das frequências e é dada em PPM (partes por milhão). A variável *stability* foi extraída do banco de dados utilizando a consulta:

```
select stability from server where ((status >> 8) & 63) != 0;
```

A Tabela 4.14 sumariza os valores da estabilidade dos relógios dos sistemas. O menor valor encontrado no conjunto de dados é zero, indicando sistemas com oscilador bastante estável. O maior valor encontrado para a estabilidade está na ordem de milhares de PPM. Apesar disso, 97% dos dados tem estabilidade inferior a 33 PPM. A grande variabilidade nos dados eleva o COV para 80 PPM, mas 50% dos dados têm estabilidade inferior a 0,015 PPM. Para 75% dos sistemas, a estabilidade é inferior a 0,239 PPM, o que equivale a um desvio de 20 ms por dia.

O gráfico de frequência da estabilidade dos relógios dos sistemas é apresentado à esquerda da Figura 4.5. Nota-se que a classe de 0,001 a 0,01 PPM é predominante, com 36% dos dados. A classe de 0,01 a 0,1 PPM é a segunda mais frequente, com 26% dos dados. Observa-se também que os valores maiores são menos frequentes. A classe para sistemas com estabilidade maior que 100 PPM contém apenas 1,5% dos dados.

As curvas da frequência cumulativa da estabilidade medida nos servidores são apre-

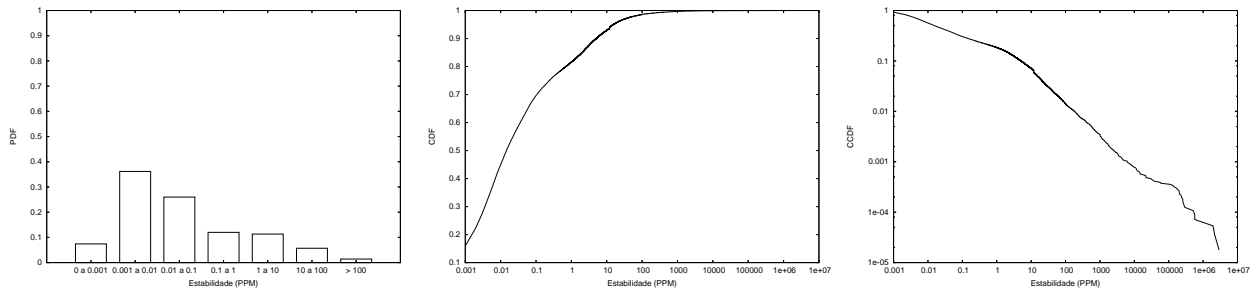


Figura 4.5: PDF, CDF e CCDF da estabilidade dos relógios

sentadas no centro da Figura 4.5. Cerca de 93% dos dados indica estabilidade de até 10 PPM. Na curva do complemento da frequência cumulativa, apresentada à direita da Figura 4.5, nota-se a existência de sistemas com osciladores problemáticos. Cerca de 3% dos *hosts* apresentam, provavelmente, problemas em seus osciladores, o que eleva a média para 228 PPM. Eliminando esses dados, a média cai para 2,32 PPM, com desvio padrão igual a 9,6 PPM. Esta média equivale a uma alteração de cerca de 200 ms por dia.

4.4 Correlação das Variáveis NTP

Como foi visto no Capítulo 2, o processo de sincronização utilizado pelo NTP depende diretamente dos valores de RTT. Quanto menor for o RTT melhor será o *offset*. Para verificarmos o efeito do atraso na rede com o *offset*, fizemos a correlação das variáveis do sistema utilizando um gráfico do tipo *box and whisker*. Os dados neste tipo de gráfico são representados com uma barra e duas linhas. Este tipo de gráfico permite mostrar de maneira simples cinco valores estatísticos dos dados. O ponto na extremidade da linha, abaixo da barra, representa o menor valor; a parte inferior da barra representa o primeiro quartil; o ponto no meio da barra representa a mediana; o ponto acima da barra representa o terceiro quartil. Por fim, o ponto na extremidade da linha, acima da barra, representa o maior valor.

No gráfico da Figura 4.6 observamos a correlação da variável atraso para a raiz com o *offset* do sistema. Neste gráfico os maiores valores não estão sendo apresentados para melhorar a visualização dos demais pontos. Os dados estão separados por classes de atraso para a raiz. Observamos que as classes para valores maiores de atraso apresentam

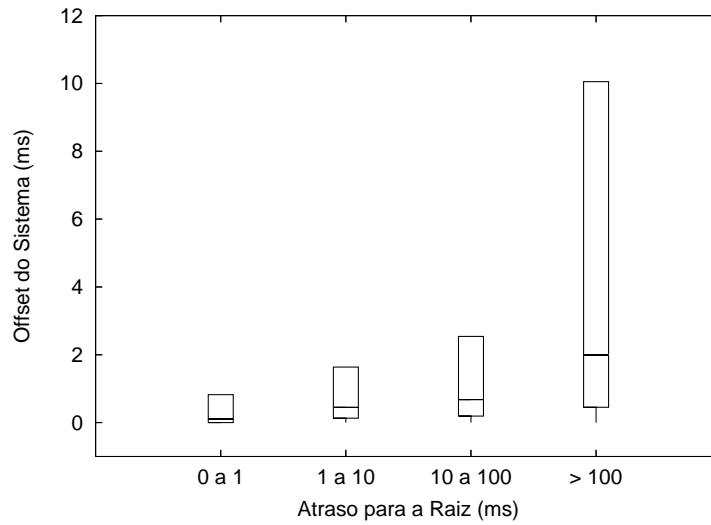


Figura 4.6: Correlação das variáveis de sistema: atraso para a raiz e *offset*.

os maiores valores de *offset*. O espaçamento entre os extremos da barra é uma medida da variabilidade dos dados. A variabilidade dos dados é proporcional ao atraso para a raiz. Os melhores *offsets* são observados na primeira classe, com os valores mínimos de atraso.

Outras variáveis de sistema que podem ser correlacionadas são a dispersão para a raiz e a estabilidade do relógio. A dispersão acumulada representa o erro máximo em relação a fonte de referência na raiz da subárvore de sincronização. A estabilidade do relógio do sistema indica a qualidade do oscilador do relógio local. No gráfico *box and whisker* da Figura 4.7 observamos que as classes com os menores valores de estabilidade apresentam também valores menores e menos variáveis para a dispersão acumulada. Neste gráfico os maiores valores não estão sendo apresentados para melhorar a visualização dos demais pontos. É possível notar que a classe com valores de estabilidade entre 1 e 10 PPM apresenta valores para a dispersão com menor variabilidade que a classe com valores de estabilidade entre 0,1 a 1 PPM. Ainda assim, percebemos maior variabilidade da dispersão para a raiz em classes com valores de estabilidade mais elevados.

4.5 Síntese do Capítulo

Os resultados apresentados neste capítulo demonstram que a rede de sincronização do NTP cresceu significativamente ao longo do período avaliado. No entanto, o acesso aos

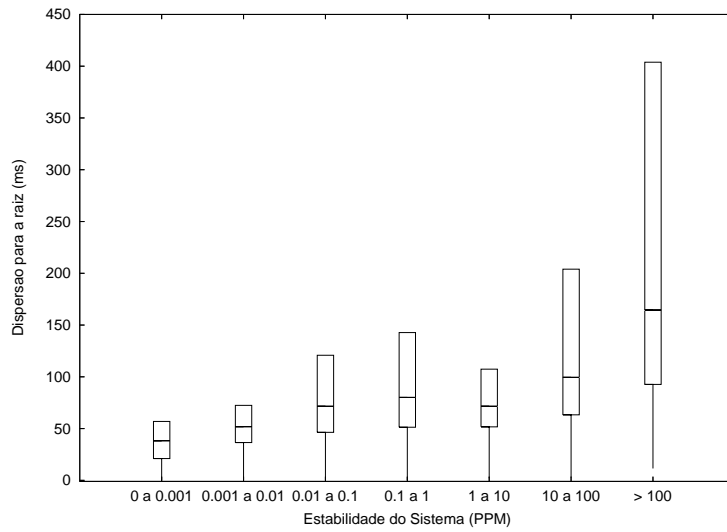


Figura 4.7: Correlação das variáveis de sistema: estabilidade e dispersão para a raiz.

computadores foi dificultado principalmente por razões de segurança, o que impede uma caracterização mais ampla da rede. A hierarquia da rede NTP foi quantificada em relação aos níveis da rede, identificados pelos *strata*. A fração de servidores primários (stratum 1) diminuiu, enquanto a fração de *hosts* stratum 3 aumentou. Isto indica que há uma procura maior por sincronização de relógios, porém há menos fontes primárias para servir a informação de tempo. Assim, a distribuição da informação de tempo depende de menos servidores, mais sobrecarregados, o que pode comprometer a qualidade e a disseminação de tempo na rede de sincronização.

Os resultados indicam melhoria em todas as métricas relacionadas à qualidade da sincronização: *offset*, RTT para o par e para a raiz, dispersão e estabilidade. O processo de sincronização utilizado pelo NTP depende diretamente de valores baixos de RTT para a obtenção de valores mais exatos de *offset*. Como o RTT diminuiu ao longo dos 15 anos, refletindo a melhoria da infra-estrutura da Internet observada no período, o *offset* também diminuiu. Melhorias nas diversas versões do protocolo e na qualidade dos osciladores dos relógios podem também ser apontadas como fatores que possivelmente contribuíram para a maior exatidão do tempo na rede. Ainda assim, a grande variabilidade observada nos valores do *offset*, medida pelo coeficiente de variação, demonstra que há computadores com estampas de tempo inexatas ou cujos relógios apresentam problemas.

Em resumo, observamos uma evolução significativa na sincronização de relógios efetu-

ada pelo NTP. No entanto, a pequena melhoria observada na métrica dispersão pode ser um indício de que a sincronização realizada pelo NTP pode estar chegando a um limite, a partir do qual melhorias adicionais podem ser mais difíceis de serem obtidas.

Capítulo 5

Aspectos Topológicos da Rede NTP

A rede formada pelos *hosts* NTP identificados na coleta de dados forma uma topologia complexa que pode ser analisada em vários aspectos. Este capítulo descreve os resultados dessa análise. As seções deste capítulo estão organizadas como descrito a seguir. A Seção 5.1 mostra a visualização da rede de sincronização do NTP; a Seção 5.2 apresenta informações sobre a localização geográfica dos *hosts* NTP; a Seção 5.3 apresenta informações dos sistemas autônomos descobertos a partir do endereçamento IP dos *hosts*; Seção 5.4 discute outras características da topologia da rede NTP. Por fim, a Seção 5.5 faz a síntese do capítulo.

5.1 Visualização da Rede NTP

A principal característica da topologia da rede NTP é a hierarquia dos *hosts*, definida pelo *stratum*. Para permitir a visualização da rede de sincronização do NTP, nos diversos níveis de *stratum*, algumas subárvores de sincronização foram plotadas utilizando a ferramenta *Otter* [30]. Esta ferramenta permite a representação de redes grandes, representadas por conjuntos de nodos e arestas.

A Figura 5.1 apresenta a visualização da rede NTP para os *hosts stratum 2* e seus pares de sincronização, os servidores de tempo *stratum 1*. As raízes das subárvores de sincronização são compostas pelos nodos *stratum 1*. O posicionamento dos nodos *stratum 1* utiliza o critério circular da ferramenta *Otter*; por isso, nota-se um círculo central com

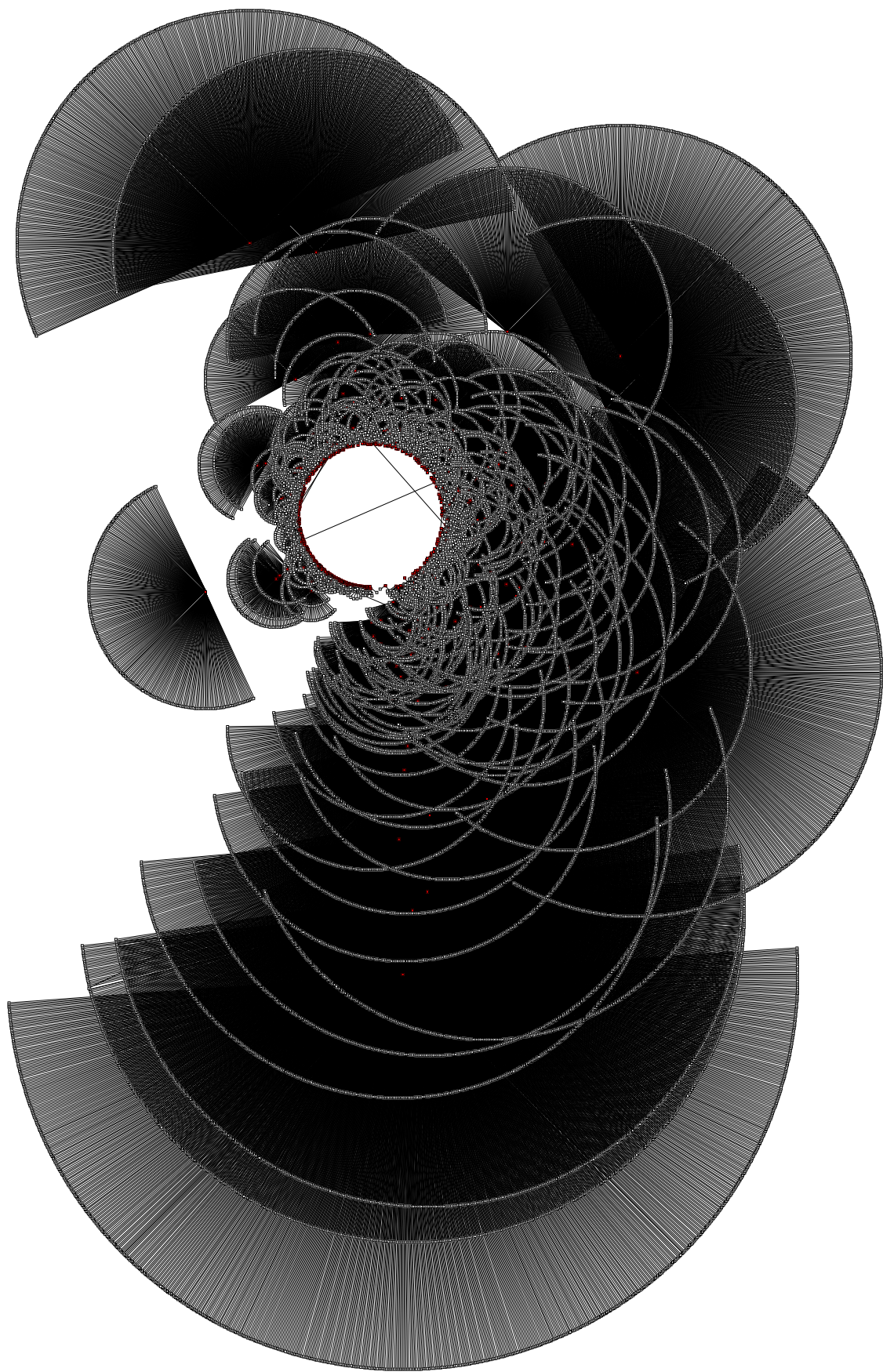


Figura 5.1: Visualização das associações da rede NTP entre os *hosts stratum 2* e seus pares de sincronização *stratum 1*.

uma grande concentração de nodos raízes. Podemos observar que os nodos *stratum* 1 são *hubs*, e possuem grande quantidade de arestas, formando diversos arcos. Quanto maior o raio do arco, maior a quantidade de pares associados ao nodo. Esses *hubs* com grande concentração de associações são apresentados pela ferramenta preferencialmente na periferia da figura.

A Figura 5.2 representa a visualização da rede NTP para os hosts *stratum* 3 e o seus pares de sincronização associados, os servidores de tempo *stratum* 2. Podemos observar que alguns nodos *stratum* 2, classificados como *hubs*, formam diversos arcos nos extremos da figura. A Figura 5.3 apresenta a visualização dos *hosts* NTP com seus pares de sincronização do nível *stratum* 3 até o nível *stratum* 16 da hierarquia. Notamos que alguns nodos *hubs* possuem servidores associados que fornecem tempo para outros *hosts* que estão em um nível abaixo na hierarquia NTP, formando pequenos arcos nas extremidades da figura.

5.2 Localização Geográfica dos Nodos

A localização geográfica dos *hosts* encontrados na coleta de dados fornece informação sobre a dispersão geográfica dos *hosts*, a abrangência da coleta e como os pares estão relacionados, levando-se em consideração a distância geográfica. Para fazer essa análise foi utilizada a ferramenta *GeoIP* [8], que retorna informações geográficas de um endereço IP. Muitas ferramentas que fazem esse tipo de consulta são comerciais. O *GeoIP*, apesar de ter uso comercial, possui um banco de dados gratuito, atualizado até março de 2002. As ferramentas desse gênero buscam descobrir com precisão a localização de um endereço IP. A precisão é superior a uma consulta feita em servidores *Whois*. Apesar da base de dados desatualizada, podemos obter, com boa precisão algumas informações geográficas básicas, como o país, território ou colônia onde se encontra o endereço.

Para cada endereço IP distinto encontrado na coleta, foi pesquisado o país onde estava localizado o *host*. Para alguns endereços, a ferramenta *GeoIP* não determinou o país onde o endereço se encontrava. Em geral, isso ocorreu para endereços com *hosts* em redes que utilizam satélites e endereços de redes privadas. Em alguns casos, a região geográfica

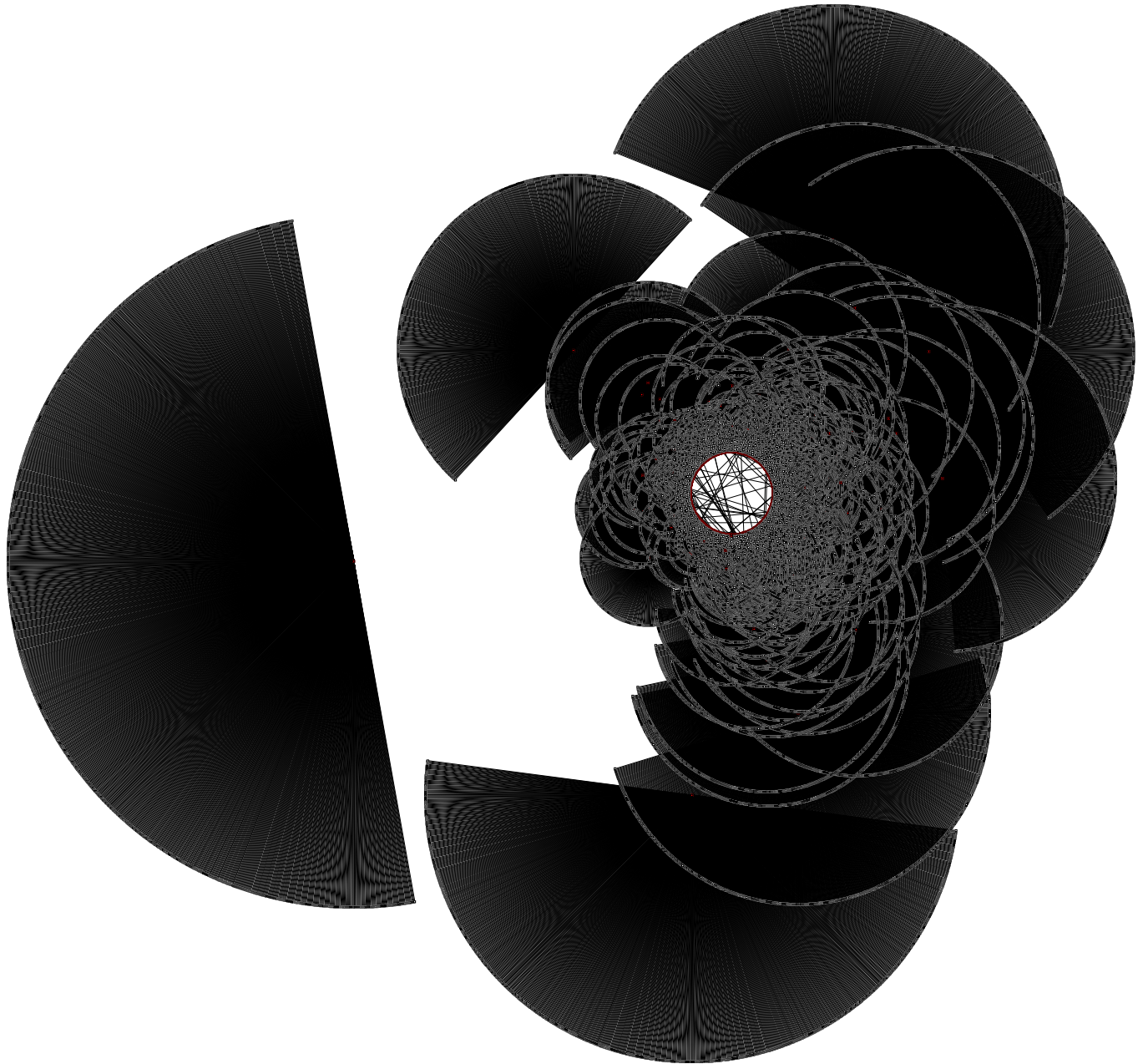


Figura 5.2: Visualização das associações da rede NTP entre os *hosts stratum 3* e seus pares de sincronização *stratum 2*.

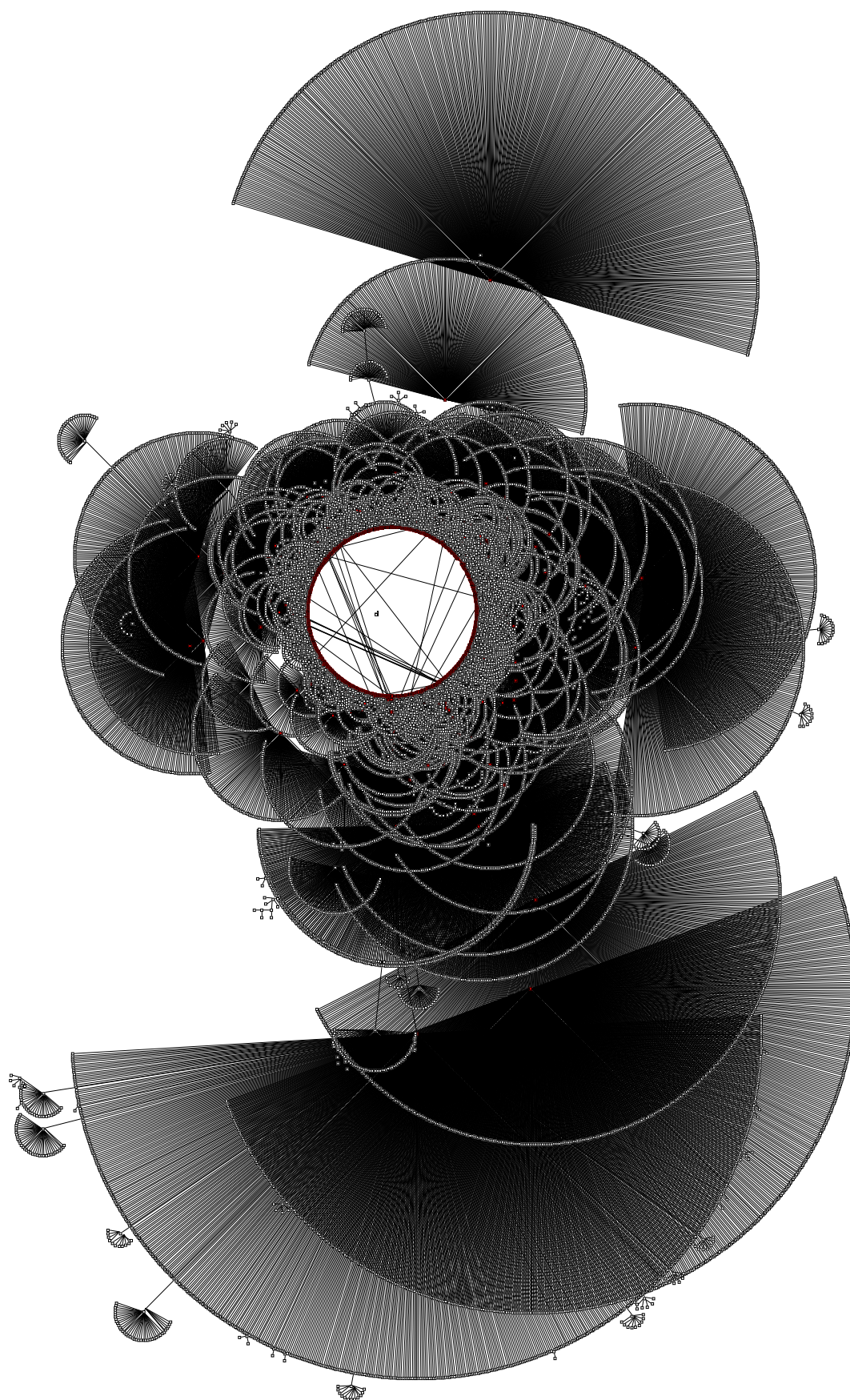


Figura 5.3: Visualização das associações pares de sincronização da rede NTP entre os *hosts* do nível *stratum* 3 até o nível *stratum* 16.

retornada não era um país, mas uma colônia de outro país ou um país que deixou de existir, como é o caso da Iugoslávia.

Os *hosts* dos endereços IP localizados na coleta de dados estão distribuídos em 158 países. Os 533 servidores *stratum* 1 estão em 47 países, enquanto os servidores *stratum* 2 e 3 estão localizados em mais de 150 países. A Tabela 5.1 mostra a quantidade de *hosts* em cada país, com um total de 147.251 *hosts*. O país que mais mantém *hosts* são os Estados Unidos, com mais de 50.000, seguido do Japão, com mais de 12.000, Inglaterra, com mais de 10.000, Alemanha, com mais de 8.000 e Rússia, com mais de 5.000. O Brasil aparece na vigésima primeira posição mundial e em primeiro lugar na América Latina, com 1361 *hosts*. O México é o segundo na América Latina, com 251 *hosts*, seguido da Argentina, com 229 *hosts*.

Do total de *hosts* identificados, 137.099 possuem um par de sincronização escolhido pelo protocolo para fornecer o tempo. A quantidade de servidores de tempo distintos para esses *hosts* é de apenas 8.152. A Tabela 5.2 mostra a localização desses servidores de tempo ao redor do mundo. O país com a maior quantidade de pares de sincronização são os Estados Unidos, com 2.707 do total (quase 1/3 do total), seguido do Japão com 708. Endereços classificados como Não Definido somam 505 servidores. O Brasil aparece com 83 servidores de tempo.

Os 8.152 servidores que atuam como par de sincronização mantêm diversas associações ativas com outros *hosts*, em uma relação cliente/servidor. A Tabela 5.3 apresenta uma lista parcial da quantidade dessas relações agrupadas por país. O resultado revela que em 90% dos casos um cliente escolhe um par de sincronização que está em seu próprio país. A tupla (país do cliente, país do servidor) com o maior número de associações é (Estados Unidos, Estados Unidos), com 49.932, seguida da tupla (Japão, Japão), com 11.684 e (Inglaterra, Inglaterra) com 8.201. A primeira ocorrência de tuplas entre países diferentes ocorre com (Rússia, Estados Unidos), com 537 associações, seguido da tupla (Canadá, Estados Unidos), com 424. Quando contabilizamos as associações apenas para pares de sincronização *stratum* 1, 70% ocorre dentro do mesmo país. Apesar da baixa quantidade de servidores *stratum* 1, muitas associações com esses pares de sincronização

País	Quantidade	País	Quantidade	País	Quantidade
Estados Unidos	55.476	Chile	93	Zâmbia	5
Japão	12.435	Turquia	84	Não Definido	4
Inglaterra	10.789	Emirados Árabes	75	Bahrein	4
Alemanha	8.255	Cazaquistão	64	Belize	4
Rússia	5.616	Arábia Saudita	62	Bolívia	4
Canadá	4.846	Equador	61	Kuwait	4
França	4.490	Europa	55	Maurício	4
Austrália	3.276	Egito	50	Uganda	4
Coréia do Sul	2.919	Croácia	46	Vanuatu	4
Itália	2.781	Venezuela	45	Albânia	3
Holanda	2.689	Bósnia-Herzegovina	43	Ilhas Virgens	3
Noruega	2.295	Irã	39	Iraque	3
Dinamarca	2.018	Azerbaijão	34	Líbano	3
Finlândia	1.887	Nicarágua	33	Madagascar	3
China	1.866	Luxemburgo	28	Namíbia	3
Suíça	1.734	Moldova	27	Antilhas Holandesas	2
Ucrânia	1.560	Bielorrússia	26	Aruba	2
Polônia	1.484	Papua Nova Guiné	24	Butão	2
Taiwan	1.479	Porto Rico	20	Fiji	2
Suécia	1.476	Uruguai	20	Gibraltar	2
Brasil	1.361	Gana	19	Iêmen	2
Tailândia	1.254	Macau	17	Ilhas Caymans	2
República Tcheca	1.248	El Salvador	15	Moçambique	2
Hong Kong	1.188	Samoa	15	Palestina	2
África do Sul	1.170	Guatemala	13	Qatar	2
Espanha	1.108	Ilhas Faroé	13	Samoa Americana	2
Áustria	1.009	Sri Lanka	13	San Marino	2
Malásia	920	Geórgia	12	Seicheles	2
Hungria	913	Bermuda	11	Somália	2
Eslovênia	454	Costa Rica	11	Sudão	2
Estônia	416	Nigéria	11	Tajiquistão	2
Eslováquia	412	Paquistão	11	Trinidade e Tobago	2
Bélgica	408	Quirguistão	11	Usbequistão	2
Lituânia	405	Jordânia	10	Vaticano	2
Portugal	392	Barbados	9	Algéria	1
Nova Zelândia	378	Cuba	9	Antígua e Barbuda	1
Irlanda	318	Macedônia	9	Bahamas	1
México	251	Bangladesh	8	Benin	1
Bulgária	245	Chipre	8	Brunei	1
Indonésia	243	Ilhas Cook	8	Camboja	1
Letônia	233	Paraguai	8	Etiópia	1
Argentina	229	Tanzânia	8	Gabão	1
Grécia	225	Peru	7	Guam	1
Filipinas	213	Quênia	7	Ilhas Salomão	1
Provedor Satélite	208	Angola	6	Listenstaine	1
Índia	205	Honduras	6	Mali	1
Singapura	180	Laos	6	Marrocos	1
Romênia	177	Malta	6	Mônaco	1
Islândia	128	Panamá	6	Nepal	1
Israel	117	Armênia	5	Omã	1
Colômbia	116	Botsuana	5	República Dominicana	1
Vietnã	106	Jamaica	5	Santa Lúcia	1
Iugoslávia	97	Nova Caledônia	5	Zimbabwe	1

Tabela 5.1: Localização dos *hosts* NTP no mundo.

País	Quantidade
Estados Unidos	2.707
Japão	708
Não Definido	505
Inglaterra	472
Alemanha	453
Rússia	412
Austrália	238
Canadá	235
Holanda	231
França	159
Suíça	136
Ucrânia	134
Suécia	122
Finlândia	120
Polônia	102
Espanha	95
Itália	92
Noruega	88
Brasil	83
República Tcheca	80
África do Sul	57
Taiwan	55
Dinamarca	49
Áustria	46
Tailândia	43
China	42
Nova Zelândia	42
Estônia	40
Portugal	40
Coréia do Sul	36

País	Quantidade
Cazaquistão	3
Emirados Árabes	3
Moldova	3
Porto Rico	3
Bangladesh	2
Bielorrússia	2
El Salvador	2
Gana	2
Guatemala	2
Ilhas Cook	2
Ilhas Faroé	2
Irã	2
Nicarágua	2
Quirguistão	2
San Marino	2
Angola	1
Bósnia-Herzegovina	1
Guam	1
Ilhas Virgens	1
Macau	1
Macedônia	1
Nova Caledônia	1
Paquistão	1
Peru	1
Samoa	1
Samoa Americana	1
Sri Lanka	1
Vaticano	1
Venezuela	1
Vietnã	1

País	Quantidade
Eslováquia	36
Bélgica	35
Eslovênia	34
Irlanda	33
Hong Kong	32
Hungria	26
Romênia	26
Grécia	20
Filipinas	19
Indonésia	18
Bulgária	17
Lituânia	17
Malásia	17
Singapura	16
México	13
Chile	12
Islândia	12
Argentina	11
Israel	11
Croácia	10
Índia	9
Letônia	9
Turquia	9
Provedor Satélite	8
Egito	7
Iugoslávia	6
Arábia Saudita	5
Luxemburgo	5
Equador	4
Papua Nova Guiné	4

Tabela 5.2: Localização dos servidores de tempo no mundo.

(País do Cliente, País do Servidor)	Quantidade
(Estados Unidos, Estados Unidos)	49.932
(Japão, Japão)	11.684
(Inglaterra, Inglaterra)	8.201
(Alemanha, Alemanha)	7.206
(França, França)	3.942
(Canadá, Canadá)	3.657
(Rússia, Rússia)	3.656
(Austrália, Austrália)	2.885
(Itália, Itália)	2.124
(Coréia do Sul, Coreia do Sul)	2.108
(Noruega, Noruega)	2.069
(Holanda, Holanda)	2.022
(Dinamarca, Dinamarca)	1.842
(Finlândia, Finlândia)	1.402
(Suíça, Suíça)	1.304
(Suécia, Suécia)	1.244
(Taiwan, Taiwan)	1.220
(China, China)	1.162
(Polônia, Polônia)	1.139
(Tailândia, Tailândia)	1.117
(Hong Kong, Hong Kong)	986
(República Tcheca, República Tcheca)	986
(África do Sul, África do Sul)	941
(Brasil, Brasil)	930
(Ucrânia, Ucrânia)	868
(Espanha, Espanha)	797
(Hungria, Hungria)	796
(Áustria, Áustria)	684
(Malásia, Malásia)	612
(Rússia, Estados Unidos)	537

(País do Cliente, País do Servidor)	Quantidade
(Canadá, Estados Unidos)	424
(Eslovênia, Eslovênia)	410
(Inglaterra, Estados Unidos)	393
(Estônia, Estônia)	365
(Lituânia, Lituânia)	337
(Estados Unidos, Canadá)	311
(Eslováquia, Eslováquia)	285
(Ucrânia, Estados Unidos)	274
(Nova Zelândia, Nova Zelândia)	269
(Bélgica, Bélgica)	258
(Estados Unidos, Inglaterra)	227
(Holanda, Alemanha)	200
(Alemanha, Estados Unidos)	192
(China, Estados Unidos)	179
(Canadá, Espanha)	171
(Letônia, Letônia)	168
(Portugal, Portugal)	166
(Estados Unidos, Alemanha)	165
(Irlanda, Irlanda)	164
(Rússia, Suécia)	164
(Japão, Estados Unidos)	151
(Alemanha, Inglaterra)	149
(Grécia, Grécia)	149
(Indonésia, Indonésia)	147
(Estados Unidos, Holanda)	145
(Rússia, Alemanha)	144
(Bulgária, Bulgária)	144
(México, Estados Unidos)	142
(Austrália, Estados Unidos)	141
(Argentina, Argentina)	132

Tabela 5.3: Distribuição dos pares de sincronização cliente/servidor entre os países (lista parcial).

estão geograficamente próximas.

A Tabela 5.4 apresenta uma lista parcial da quantidade de países clientes por país de localização do par de sincronização. Os Estados Unidos aparece em primeira posição, fornecendo tempo para outros 126 países, seguido da Inglaterra, com 70 vizinhos clientes e da Alemanha com 63. O Brasil fornece tempo para 16 países.

A Tabela 5.5 lista os países dos *hosts* clientes do Brasil. Na primeira posição aparece o próprio Brasil, com 930 associações, confirmando o princípio que os clientes escolhem geralmente servidores próximos, em seu próprio país. Outros clientes são os Estados Unidos e a Inglaterra, países que não são geograficamente próximos. Ainda assim, é possível verificar a presença de países sul-americanos, como Colômbia, Argentina e Paraguai.

A rede NTP constrói alguns *hubs* nos diversos países e esses servidores são na maioria dos casos, pares de sincronização de *hosts* localizados no mesmo país, fazendo dos nodos da rede NTP, em geral, geograficamente próximos.

País servidor de tempo	Países Clientes
Estados Unidos	126
Inglaterra	70
Alemanha	63
França	59
Suécia	57
Holanda	54
Canadá	50
Austrália	49
Suíça	40
Dinamarca	40
Bélgica	36
Polônia	35
Itália	34
Noruega	30
Espanha	29
Rússia	29
Eslovênia	28
Japão	27
Finlândia	25
Hong Kong	25

País servidor de tempo	Países Clientes
República Tcheca	23
Lituânia	23
Irlanda	21
Áustria	20
Hungria	18
Singapura	17
Brasil	16
Nova Zelândia	13
Argentina	13
Romênia	13
Croácia	12
Turquia	10
Grécia	10
África do Sul	10
Coréia do Sul	9
Malásia	9
Chile	9
México	9
Tailândia	8
Luxemburgo	8

Tabela 5.4: Países servidores de tempo e número de clientes (lista parcial).

País	Quantidade de <i>Hosts</i> Clientes
Brasil	930
Estados Unidos	25
Inglaterra	3
Colômbia	2
Coréia do Sul	2
México	2
Nova Zelândia	2
Provedor Satélite	2
Suécia	2
Uruguai	2
Argentina	1
Austrália	1
China	1
Indonésia	1
Nicarágua	1
Paraguai	1

Tabela 5.5: Países com *hosts* clientes do Brasil.

5.3 Identificação dos Sistemas Autônomos

Além da informação geográfica de cada endereço IP identificado na coleta de dados, outra informação que pode ser analisada a partir do endereçamento IP é a obtida através do protocolo - *Border Gateway Protocol* (BGP). O protocolo BGP permite a troca de informação de roteamento entre os diversos domínios existentes na Internet. Cada domínio é conhecido como um Sistema Autônomo (AS - do inglês *Autonomous System*) e é identificado com um número distinto chamado *Autonomous System Number* (ASN). Em geral, um endereço IP pertence a um único AS. Pela tabela de roteamento BGP é possível descobrir que AS está anunciando um determinado endereçamento IP. Consultando o cadastro de ASes disponível no sistema *Whois*, podemos obter informações do provedor Internet responsável pelo endereçamento IP.

Para cada endereço IP distinto encontrado na coleta pesquisamos a que AS pertencia esse endereço. Essa consulta foi feita utilizando-se tabelas de roteamento BGP de acesso público, disponíveis em visualizadores de rotas *route-views*, do projeto da Universidade de Oregon [29]. Esses visualizadores de rotas obtêm informações de diversos backbones Internet, localizados em diversas partes do mundo. Para cada identificador de AS obtido foram coletadas as informações do provedor consultando a base de dados *Whois*.

Os endereços IP de *hosts* identificados na coleta de dados estão distribuídos em 5.861 sistemas autônomos. A Tabela 5.6 apresenta uma lista parcial do número de *hosts* em cada AS. Os sistemas autônomos que mais mantêm *hosts* NTP estão na comunidade européia, nos Estados Unidos e no Japão. O provedor brasileiro com maior quantidade de servidores é o ASN 4230 (EMBRATEL).

A quantidade de pares de sincronização distintos para os *hosts* NTP é 8.152. A Tabela 5.7 mostra a quais sistemas autônomos pertencem esses pares de sincronização. Os sistemas autônomos que mais mantêm pares de sincronização estão no Estados Unidos, Japão e comunidade européia. Nessa tabela observamos a presença de muitos institutos que promovem atividades de pesquisa, como as redes de educação e pesquisa da Suíça, Austrália, Noruega, França, Espanha, Itália, Estados Unidos, Brasil e outras. No Brasil, o provedor com a maior quantidade de pares de sincronização é a Rede Nacional de Ensino

ASN	Nome do AS	Localização	Quantidade
786	JANET The JANET IP Service	Europa	2.230
680	DFN-IP service G-WiN	Europa	1.958
224	UNINETT UNINETT, The Norwegian University Research Network	Europa	1.714
4589	EASYNET Easynet Group Plc	Europa	1.625
209	ASN-QWEST - Qwest	Estados Unidos	1.547
2907	ERX-SINET-AS National Center for Science Information Systems	Japão	1.529
59	WISC-MADISON-AS - University of Wisconsin Madison	Estados Unidos	1.509
18127	JPNIC-ASBLOCK-AP JPNIC	Japão	1.339
7018	ATT-INTERNET4 - ATT WorldNet Services	Estados Unidos	1.202
15557	LDCOMNET NEUF CEGETEL (formerly LDCOM NETWORKS)	França	1.197
159	SONNET-AS - The Ohio State University	Estados Unidos	1.108
9	CMU-ROUTER - Carnegie Mellon University	Estados Unidos	1.087
137	ASGARR GARR Italian academic and research network	Europa	983
13237	LAMBDANET-AS European Backbone of LambdaNet	Europa	957
2914	NTT-COMMUNICATIONS-2914 - NTT America, Inc.	Estados Unidos	939
3	MIT-GATEWAYS - Massachusetts Institute of Technology	Estados Unidos	909
7132	SBIS-AS - SBC Internet Services	Estados Unidos	905
7575	AARNET-AS-AP Australian Academic and Research Network (AARNet)	Austrália	867
4134	CHINANET-BACKBONE No.31,Jin-rong Street	China	805
701	ALTERNET-AS - UUNET Technologies, Inc.	Estados Unidos	783
559	SWITCH SWITCH, Swiss Education and Research Network	Suíça	779
5483	HTC-AS Hungarian Telecom	Europa	753
4716	POWEREDCOM KDDI Corporation	Japão	727
3741	IS	África do Sul	722
6106	UCR-EDU - University of California, Riverside	Estados Unidos	715
3313	INET-AS I.NET S.p.A.	Europa	713
17707	EDGE-JP-AP AS for DATAHOTEL, which is one of iDC in Japan,	Japão	698
3786	ERX-DACOMNET DACOM Corporation	Coreia do Sul	698
15290	ALLST-15290 - Allstream Corp. Corporation Allstream	Canadá	695
5462	CABLEINET Telewest Broadband	Inglaterra	684
3356	LEVEL3 Level 3 Communications	Estados Unidos	681
1103	SURFNET-NL SURFnet, The Netherlands	Europa	657
2042	ERX-JARING Malaysian institute of Microelectronic Systems (MIMOS)	Malásia	652
2152	CSUNET-NW - California State University Network	Estados Unidos	638
18	UTEXAS - University of Texas at Austin	Estados Unidos	628
2381	WISCNET1-AS - University of Wisconsin-Madison	Estados Unidos	623
5413	AS5413 PIPEX Communications	Inglaterra	618
3320	DTAG Deutsche Telekom AG	Alemanha	615
25	UCB - University of California at Berkeley	Estados Unidos	610
2828	XO-AS15 - XO Communications	Estados Unidos	604
19548	ADELPHIA-AS2 - Adelphia	Estados Unidos	600
3561	SAVVIS - Savvis	Estados Unidos	596
29484	RUB-AS Ruhr-Universitaet Bochum	Alemanha	585
12670	Completel Autonomous System in France	França	577
766	REDIRIS RedIRIS Autonomous System	Europa	570
8220	COLT COLT Telecommunications	Inglaterra	566
237	MERIT-AS-14 - Merit Network Inc.	Estados Unidos	566
46	RUTGERS - Rutgers University	Estados Unidos	556
6785	CYBERCITY Cybercity AS	Dinamarca	552
3388	UNM-AS - University of New Mexico	Estados Unidos	545
2500	JPNIC-ASBLOCK-AP JPNIC	Japão	511
790	EUNETFI EUnet Finland	Europa	509
553	BELWUE Landeshochschulnetz Baden-Wuerttemberg (BelWue)	Europa	499
2852	CESNET2 Czech National Research	Europa	496
4725	ODN JAPAN TELECOM CO.,LTD.	Japão	485
8111	DALUNIV - Dalhousie University	Canadá	484
9370	SAKURA-B SAKURA Internet Inc.	Japão	466
702	AS702 MCI EMEA - Commercial IP service provider in Europe	Estados Unidos	466
195	SDSC-AS - San Diego Supercomputer Center	Estados Unidos	455
14	COLUMBIA-GW - Columbia University	Estados Unidos	455
8342	RTCOMM-AS RTComm.RU Autonomous System	Rússia	448
2637	GEORGIA-TECH - Georgia Institute of Technology	Estados Unidos	440
9488	SNU-AS-KR Seoul National University	Coreia do Sul	440
4750	LOXLEY-TH Loxley Information Co., Ltd.	Tailândia	436
4230	EMBRATEL	Brasil	420

Tabela 5.6: Quantidade de *Hosts* NTP em cada sistema autônomo (lista parcial).

e Pesquisa (RNP), que aparece fornecendo tempo através de 16 *hosts* NTP.

Os 8.152 servidores que atuam como pares de sincronização mantêm diversas associações ativas com outros *hosts*, em uma relação cliente/servidor. A Tabela 5.8 mostra uma lista parcial da quantidade dessas relações agrupadas por AS. O resultado revela que em 68% dos casos um cliente escolhe um par de sincronização que está em seu domínio, ou seja, que pertence ao mesmo AS. Este fato pode estar relacionado com a recomendação de configuração do NTP, para o uso de servidores locais para consultas de tempo e, desta forma, evitar tráfego desnecessário na Internet. A tupla (AS do cliente, AS do servidor) com o maior número de associações é entre o ASN 786. Quando contabilizamos as associações apenas para pares de sincronização *stratum* 1, 32% ocorrem dentro o mesmo AS.

A Tabela 5.9 mostra todos os sistemas autônomos que foram identificados como brasileiros e que mantêm algum par de sincronização. Observamos a presença de muitos institutos que promovem a pesquisa no Brasil, como a Rede Nacional de Ensino e Pesquisa, a Fundação de Amparo à Pesquisa do Estado de São Paulo, a Universidade de São Paulo, a Universidade Federal do Paraná e outros sistemas autônomos.

5.4 Outras Características Topológicas

A topologia da rede NTP pode ainda ser caracterizada de três maneiras. A primeira é a quantidade de associações configurada em um servidor. O NTP utiliza um algoritmo para rejeitar uma origem de tempo ruim e, dinamicamente, seleciona o melhor par de sincronização em um conjunto de pares disponíveis. Manter mais de uma associação, além de garantir redundância, possibilita que o NTP utilize métodos estatísticos para escolher a origem do tempo com maior confiabilidade.

Para a análise da distribuição do número de associações por servidor, a consulta abaixo foi executada no banco de dados, para contabilizar a quantidade de endereços de origem distintos nas associações de um servidor, considerando servidores com *stratum* 1, 2 e 3.

```
select count(distinct p.srcadr) from server s, peer p where s.server = p.server
and s.stratum = 1,2,3 group by s.server;
```

ASN	Nome do AS	Localização	Quantidade
2914	NTT-COMMUNICATIONS-2914 - NTT America, Inc.	Estados Unidos	182
786	JANET The JANET IP Service	Europa	175
2907	ERX-SINET-AS National Center for Science Information Systems	Japão	129
680	DFN-IP service G-WiN	Europa	127
559	SWITCH SWITCH, Swiss Education and Research Network	Suíça	73
1103	SURFNET-NL SURFnet, The Netherlands	Europa	67
701	ALTERNET-AS - UUNET Technologies, Inc.	Estados Unidos	67
209	ASN-QWEST - Qwest	Estados Unidos	64
7575	AARNET-AS-AP Australian Academic and Research Network (AARNet)	Austrália	57
224	UNINETT UNINETT, The Norwegian University Research Network	Europa	56
3356	LEVEL3 Level 3 Communications	Estados Unidos	55
2200	FR-RENATER Reseau National de telecommunications pour la Technologie	Europa	54
766	REDIRIS RedIRIS Autonomous System	Europa	48
137	ASGARR GARR Italian academic and research network	Europa	46
8342	RTCOMM-AS RTComm.RU Autonomous System	Rússia	43
2500	JPNIC-ASBLOCK-AP JPNIC	Japão	42
4716	POWEREDCOM KDDI Corporation	Japão	39
3	MIT-GATEWAYS - Massachusetts Institute of Technology	Estados Unidos	39
18127	JPNIC-ASBLOCK-AP JPNIC	Japão	38
23504	SPEAKEASY - Speakeasy, Inc.	Estados Unidos	38
7132	SBIS-AS - SBC Internet Services	Estados Unidos	36
5413	AS5413 PIPEX Communications	Inglaterra	34
553	BELWUE Landeshochschulnetz Baden-Wuerttemberg (BelWue)	Europa	33
1930	RCCN RCCN-NET	Europa	33
2852	CESNET2 Czech National Research	Europa	32
46	RUTGERS - Rutgers University	Estados Unidos	29
6079	RCN-AS - RCN Corporation	Estados Unidos	29
7018	ATT-INTERNET4 - ATT WorldNet Services	Estados Unidos	29
8220	COLT COLT Telecommunications	Inglaterra	28
1653	SUNET SUNET Swedish University Network	Europa	27
5388	ENERGIS-AS Energis UK	Inglaterra	27
1239	SPRINTLINK - Sprint	Estados Unidos	26
3561	SAVVIS - Savvis	Estados Unidos	25
2915	ERX-SPIN-NET ATT Jens Corporation	Japão	25
4589	EASYNET Easynet Group Plc	Europa	25
25	UCB - University of California at Berkeley	Estados Unidos	25
5617	TPNET Polish Telecoms commercial IP network	Polônia	24
174	COGENT Cogent/PSI	Estados Unidos	24
3741	IS	África do Sul	24
237	MERIT-AS-14 - Merit Network Inc.	Estados Unidos	24
271	BCNET-AS - University of British Columbia	Canadá	23
852	ASN852 - Telus Advanced Communications	Canadá	23
4713	OCN NTT Communications Corporation	Japão	23
17055	UTAH - University of Utah	Estados Unidos	23
73	WASHINGTON-AS - University of Washington	Estados Unidos	22
3265	XS4ALL-NL XS4ALL	Holanda	22
6360	UNIVHAWAII - University of Hawaii Data Video Networks	Estados Unidos	22
9	CMU-ROUTER - Carnegie Mellon University	Estados Unidos	20
22773	CCINET-2 - Cox Communications Inc.	Estados Unidos	20
790	EUNETFI EUnet Finland	Europa	19
2107	ARNES-NET ARNES	Europa	18
5515	SONERA-GLOBAL-IP Sonera Solution Autonomous System	Finlândia	18
59	WISC-MADISON-AS - University of Wisconsin Madison	Estados Unidos	17
7377	UCSD - University of California at San Diego	Estados Unidos	17
3320	DTAG Deutsche Telekom AG	Alemanha	17
3216	SOVAM-AS Golden Telecom, Moscow, Russia	Rússia	17
6517	YIPESCOM - Yipes Communications, Inc.	Estados Unidos	17
19548	ADELPHIA-AS2 - Adelphia	Estados Unidos	17
7509	HINET Hokkaido University	Japão	17
292	ESNET-WEST - Energy Sciences Network	Estados Unidos	16
1741	FUNETAS FUNET autonomous system	Europa	16
6461	MFNX MFN - Metromedia Fiber Network	Estados Unidos	16
12816	MWN-AS Leibniz-Rechenzentrum Muenchen	Alemanha	16
2828	XO-AS15 - XO Communications	Estados Unidos	16
1916	Rede Nacional de Ensino e Pesquisa	Brasil	16

Tabela 5.7: Número de pares de sincronização em cada sistema autônomo (lista parcial).

(ASN do Cliente, ASN do Servidor)	Quantidade	(ASN do Cliente, ASN do Servidor)	Quantidade
(786, 786)	2133	(7575, 7575)	788
(680, 680)	1759	(5483, 5483)	723
(224, 224)	1623	(6106, 6106)	700
(4589, 4589)	1498	(3313, 3313)	686
(59, 59)	1491	(17707, 17707)	681
(2907, 2907)	1381	(559, 559)	658
(18127, 18127)	1332	(5462, 5462)	639
(15557, 15557)	1209	(4716, 4716)	615
(159, 159)	1087	(3786, 3786)	589
(209, 209)	1062	(2381, 2381)	586
(9, 9)	1061	(1103, 1103)	584
(13237, 13237)	940	(3741, 3741)	581
(137, 137)	918	(29484, 29484)	574
(3, 3)	868	(4134, 4134)	569
(2914, 2914)	801	(2152, 2152)	566

Tabela 5.8: Relação cliente/servidor entre os sistemas autônomos (lista parcial).

ASN	Nome do AS	Quantidade
1916	Rede Nacional de Ensino e Pesquisa	16
4230	EMBRATEL	11
2715	Fundação de Amparo à Pesquisa do Estado de São Paulo	7
2716	Fundação de Amparo à Pesquisa do Estado de São Paulo	3
7465	PROCERGS - Cia. de Processamento de Dados do RGS	3
13353	Telmex do Brasil Ltda	3
17379	Intelig Telecomunicações Ltda	2
22548	Comitê Gestor da Internet no Brasil	2
7738	Telecomunicações da Bahia S.A.	2
26622	SIM Telecomunicações Ltda	2
28571	Universidade de São Paulo - USP	2
10881	PoP-PR - Universidade Federal do Paraná	2
28573	NET Serviços de Comunicação S.A.	2
11242	PoP-SC Ponto de Presença da RNP em Santa Catarina	2
11432	BRASILNET TELECOMUNICACOES LTDA	2
13878	Diveo do Brasil Telecomunicações Ltda	2
15180	Diveo do Brasil Telecomunicações Ltda	2
1251	Fundação de Amparo à Pesquisa do Estado de São Paulo	2

Tabela 5.9: Sistemas autônomos com pares de sincronização no Brasil (lista parcial).

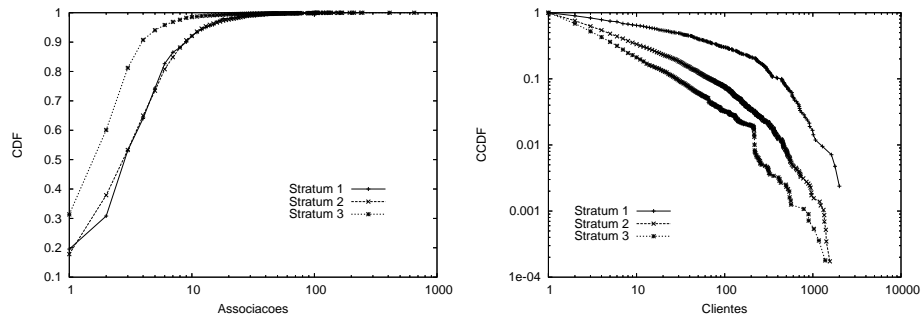


Figura 5.4: Distribuição das associações por *stratum* (esquerda) e dos clientes por *stratum* (direita).

O gráfico à esquerda da Figura 5.4 apresenta esta distribuição, com escala logarítmica no eixo x . Verifica-se que os servidores *stratum* 1 e 2 apresentam distribuições similares para a quantidade de associações. Em ambos os casos, 20% dos servidores possuem apenas uma associação e 50% possui até três associações. Para os servidores *stratum* 3, 31% tem apenas uma associação e 75% até três associações. A baixa conectividade da maioria dos servidores é surpreendente quando é verificada a quantidade de servidores *stratum* 2 e 3 disponíveis na rede.

Em todos os casos analisados há uma fração pequena porém não desprezível de servidores com muitas associações. Há servidores que mantêm dezenas, ou mesmo centenas de associações. Este quadro está de acordo com o modelo de desenvolvimento da topologia da Internet denominado “livre de escala”, que tem como critério essencial a conexão preferencial [1, 31, 5]. Este critério indica que novos nodos tendem a se conectar em vértices já extremamente conectados. Este modelo tem como consequência uma distribuição muito variável em relação aos graus dos nodos [7], como mostram os gráficos da Figura 5.4.

Uma segunda maneira de caracterizar a topologia da rede NTP é avaliar o número de pares que utilizam um determinado servidor. Para isso, foi executada a consulta mostrada abaixo no banco de dados para que fosse verificado, para cada endereço de servidor conhecido, quantas vezes este endereço apareceu como endereço de origem no conjunto de pares. Esta quantidade é denominada “cliente”, e indica quantos clientes distintos cada servidor possui.

```
select count(distinct p.server) from peer p, server s where p.srcadr = s.server
and s.stratum = 1,2,3 group by p.srcadr;
```

<i>stratum</i>	1993	1995	1999	2005
1	22,36	20,17	28,03	53,45
2	2,29	1,48	3,18	2,86
3	0,59	0,34	0,45	0,30
4	0,02	0,14	0,19	0,07

Tabela 5.10: Índice de carga por *stratum* em vários estudos.

O gráfico à direita da Figura 5.4 apresenta este resultado, e demonstra uma carga maior nos servidores *stratum* 1, seguida da carga nos servidores *strata* 2 e 3. Cerca de 50% dos servidores *stratum* 1 analisados possui pelo menos 27 clientes. Possui mais de 400 clientes cerca de 10% dos servidores *stratum* 1, 1,3% dos servidores *stratum* 2 e 0,3% dos servidores *stratum* 3. Este quadro também confirma o modelo livre de escala proposto para a Internet.

A terceira maneira de caracterizar a topologia da rede NTP é avaliar a carga dos servidores NTP. Um método simples para calcular a carga de servidores *stratum* n é dividir a quantidade de servidores *stratum* $n + 1$ pela quantidade de servidores *stratum* n . A Tabela 5.10 apresenta este resultado para os diversos estudos e mostra um aumento considerável na utilização de servidores *stratum* 1. Este resultado é explicado pela diminuição do número de servidores *stratum* 1 apesar do aumento do tamanho da rede.

5.5 Síntese do Capítulo

A visualização da rede NTP contribui para a compreensão de seu tamanho, da hierarquia e da quantidade de clientes por servidor em cada nível da rede. Os nodos muito conectados, denominados *hubs*, são facilmente visíveis na estrutura desenhada.

A distribuição geográfica dos nodos mostra que a rede está presente em mais de 150 países e regiões em todo o mundo, sendo que 47 países possuem servidores primários de tempo (*stratum* 1). A identificação dos sistemas autônomos aos quais pertencem os *hosts* da rede indica as instituições provedoras de acesso Internet aos *hosts*. Observa-se que a sincronização ocorre preferencialmente entre pares geograficamente próximos, no mesmo sistema autônomo ou no mesmo país. Dois fatores podem contribuir para

este comportamento. O primeiro é a questão de confiabilidade e segurança da origem do tempo. O segundo fator é a preferência pela configuração dos *hosts* com servidores fisicamente próximos, para evitar o uso dos backbones da Internet com este tipo de tráfego (mensagens do NTP).

A análise da topologia indica que o número de associações mantidas por servidores de tempo strata 1, 2 e 3 é pequeno. O NTP utiliza algoritmos complexos para escolher a melhor fonte de tempo dentre um conjunto de fontes disponíveis. No entanto, se há poucas fontes de tempo disponíveis, a eficiência do algoritmo pode ficar comprometida, e isto pode prejudicar a sincronização dos relógios com o tempo UTC. O algoritmo escolherá uma fonte seguindo o modelo de melhor esforço, porém não há garantia da qualidade da sincronização.

Por outro lado, o número de clientes por servidor pode ser muito grande, o que significa que tempos inexatos podem ser distribuídos largamente pela rede. O pequeno número de servidores primários demonstra que a rede está baseada em poucos servidores, que experimentam carga elevada. Um estudo da qualidade de tempo fornecida pelos servidores primários deve ser realizado.

Capítulo 6

Conclusão e Trabalhos Futuros

Sistemas distribuídos podem requerer sincronização dos relógios dos nodos do sistema com o tempo mundial para disparar e registrar a execução de tarefas e a ocorrência de eventos. A falta de sincronização pode gerar falhas na execução correta das tarefas do sistema, bem como perda de dados, problemas de segurança e de confiabilidade.

Muitas aplicações que executam na Internet podem depender da sincronização realizada pelo NTP. Essa dissertação apresentou uma caracterização da rede de sincronização do NTP. Diversas variáveis disponíveis nos *hosts* alcançados na coleta de dados foram analisadas e comparadas com estudos realizados durante os últimos 15 anos.

A análise dos resultados indica que, embora a rede tenha crescido substancialmente em termos do número de nodos ou *hosts*, o acesso a estes nodos se tornou mais difícil devido a procedimentos de segurança implementados nos sistemas, o que dificulta também a caracterização desta e de outras redes. A rede alcançada é apenas parte de uma rede muito maior.

Em relação à arquitetura do NTP, observa-se que há uma porcentagem menor de servidores *stratum* 1, e estes servidores apresentam atualmente uma carga bem maior do que a observada no passado. Nota-se também um crescimento significativo da fração de servidores *stratum* 3, o que pode ser uma evidência de procura maior por fontes de sincronização mais exatas.

Embora a arquitetura do NTP defina o conceito de *stratum* e seu papel hierárquico na

rede NTP, nenhum aspecto do protocolo interfere nas relações quantitativas observadas, tais como número de servidores por *stratum*, número de associações e número de clientes de cada servidor. Estes valores são definidos por outras forças que modelam a rede. Os resultados quantitativos da caracterização da topologia mostram grande variação na distribuição das conexões dos nodos, evidenciando características do modelo livre de escala, que é utilizado atualmente para modelar redes complexas.

A rede de sincronização do NTP evoluiu e a sincronização com o tempo UTC é melhor do que no passado. A análise dos valores de atualização dos relógios (*offset*) ao longo do tempo revela que a rede está substancialmente melhor sincronizada atualmente. O *offset* médio é de 7 ms e a mediana é de 0,7 ms. Apenas 0,2% dos *hosts* apresenta diferença superior a 128 ms em relação à hora certa. A análise do tempo de ida e volta na rede (RTT), medida fundamental nos algoritmos do NTP, revela queda significativa neste valor, resultado que contribui para o bom desempenho do protocolo.

Apesar destes progressos, a exatidão da sincronização continua na ordem de poucos milissegundos. Os relógios dos computadores são sincronizados por meio de troca de mensagens via Internet, e o RTT médio está na ordem de milissegundos. Além disso, os osciladores dos relógios podem flutuar e gerar erros da ordem de dezenas ou centenas de milissegundos por dia. Esta exatidão pode não ser suficiente em algumas situações. Em um milissegundo um computador pode executar milhares de tarefas. Se for necessário registrar o tempo real e manter a ordem relativa de diversos eventos executados em computadores diferentes, esta exatidão pode ser insuficiente.

Há várias possibilidades de trabalho futuro. A qualidade da sincronização depende da exatidão e da resolução dos relógios das fontes primárias de tempo, da troca de mensagens na rede, da existência de muitos pares de sincronização, da qualidade dos osciladores dos relógios e dos próprios algoritmos do protocolo NTP. Todos estes aspectos devem ser investigados individualmente. Os limites da sincronização provida pelo NTP devem ser investigados. É possível obter exatidão da ordem de microssegundos ou nanossegundos numa rede como a Internet?

Referências Bibliográficas

- [1] A. Barabási e R. Albert. Emergence of Scaling in Random Networks. *Science*, 286, October de 1999.
- [2] Albert-Laszlo Barabási. *Linked: The New Science of Networks*. Perseus Publishing, 2002.
- [3] G. Coulouris, J. Dollimore, e T. Kindberg. *Distributed Systems - Concepts and Design*. Addison-Wesley, 2. ed., 1994.
- [4] Bureau International des Poids et Mesures. Unit of time (second). <http://www.bipm.fr>.
- [5] S.N. Dorogovtsev e J.F.F. Mendes. *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford University Press, 2003.
- [6] P. Erdős e A. Rényi. On Random Graphs I. *Debrecen*, 1959.
- [7] Michalis Faloutsos, Petros Faloutsos, e Christos Faloutsos. On Power-law Relationships of the Internet Topology. *Proc. ACM SIGCOMM'99*, 1999.
- [8] GeoIP - IP Geolocation, 2006, <http://www.maxmind.com/geoip/>.
- [9] J. D. Guyton e M. F. Schwartz. Experiences with a survey tool for discovering Network Time Protocol servers. *Usenix Summer 1994 Technical Conference*, 1994.
- [10] Sun-Mi Jun, Dong-Hui Yu, Young-Ho Kim, e Soon-Yong Seong. A Time Synchronization Method for NTP. *Real-Time Computing Systems and Applications, 1999. RTCSA '99. Sixth International Conference on 13-15 Dec*, pp. 466–473, 1999.
- [11] J. Kohl e C. Neuman. RFC 1510: The Kerberos Network Authentication Service (V5), Setembro de 1993.
- [12] K. Marzullo e S. Owicki. Maintaining the time in a distributed system. *ACM Operating Systems Review*, 19(3):44–54, 1985.
- [13] D. L. Mills. *Computer Network Time Synchronization: the Network Time Protocol*. CRC Press, 1. ed., 2006.
- [14] D. L. Mills. RFC 891: DCN Local-Network Protocols, Dezembro de 1983.
- [15] D. L. Mills. RFC 956: Algorithms for synchronizing network clocks, Setembro de 1985.

- [16] D. L. Mills. RFC 957: Experiments in network clock synchronization, Setembro de 1985.
- [17] D. L. Mills. RFC 958: Network Time Protocol (NTP), Setembro de 1985.
- [18] D. L. Mills. RFC 1059: Network Time Protocol (version 1) specification and implementation, Julho de 1988.
- [19] D. L. Mills. RFC 1119: Network Time Protocol (version 2) specification and implementation, Setembro de 1989.
- [20] D. L. Mills. On the accuracy and stability of clocks synchronized by the Network Time Protocol in the Internet system. *ACM Computer Communication Review*, 20(1):65–75, 1990.
- [21] D. L. Mills. Internet Time Synchronizantion: the Network Time Protocol. <http://www.eecis.udel.edu/~mills/database/papers/trans.ps>, Outubro de 1991.
- [22] D. L. Mills. RFC 1305: Network Time Protocol (Version 3) Specification, Implementation and Analysis, Março de 1992.
- [23] D. L. Mills. A Brief History of NTP Time: Memoirs of an Internet Timekeeper. *Computer Communication Review*, 33(2):9–21, 2003.
- [24] D. L. Mills e P. Kamp. The Nanokernel. <http://www.eecis.udel.edu/~mills/ntp/html/kern.html>, Dezembro de 2002.
- [25] D. L. Mills, A. Thyagarajan, e B.C. Huffmann. Internet timekeeping around the globe. <http://www.cis.udel.edu/~mills/database/papers/survey5.ps>, Dezembro de 1997.
- [26] N. Minar. A Survey of the NTP Network. <http://www.media.mit.edu/~nelson/research/ntp-survey99/>, Dezembro de 1999.
- [27] P. Mockapetris. RFC 1034: Domain names - concepts and facilities, Novembro de 1987.
- [28] NTP. NTP (Network Time Protocol) Public Services Project, 2005, <http://ntp.isc.org>.
- [29] University of Oregon Route Views Project, 2006, <http://www.routeviews.org/>.
- [30] Otter - Tool for Topology Display, 2006, <http://www.caida.org/tools/visualization/otter/>.
- [31] Romualdo Pastor-Satorras e Alessandro Vespignani. *Evolution and Structure of the Internet : A Statistical Physics Approach*. Cambridge University Press, 2004.
- [32] J. Postel. RFC 781: A Specification of the Internet Protocol (IP) Timestamp Option, Maio de 1981.
- [33] J. Postel. RFC 792: Internet Control Message Protocol, Setembro de 1981.
- [34] J. Postel. RFC 768: User Datagram Protocol, Maio de 1983.

- [35] J. Postel. RFC 791: Internet Protocol, Maio de 1983.
- [36] J. Postel. RFC 867: Daytime Protocol, Maio de 1983.
- [37] J. Postel. RFC 868: Time Protocol, Maio de 1983.
- [38] Peter Rybaczky. *Expert Network Time Protocol: An Experience in Time with NTP*. Apress, 2005.
- [39] D. B. Sullivan, D. W. Allan, D. A. Howe, e F. L. Walls. Characterization of Clocks and Oscillators. *National Institute of Standards and Technology*, Technical Note 1337:121–128, 1990.
- [40] Pedro Rodrigues Torres-Júnior. Coleta de Dados da Rede NTP. <http://www.ntpsurvey.arauc.br>, 2006.
- [41] M. Tsuru, T. Takine, e Y. Oie. Estimation of clock offset from one-way delay measurement on asymmetric paths. *Applications and the Internet (SAINT) Workshops, 2002. Proceedings. 2002 Symposium on 28 Jan.-1*, pp. 126–133, 2002.
- [42] N. Vasanthavada e P. N. Marinos. Synchronization of fault-tolerant clocks in the presence of malicious failures. *IEEE Trans. Computers*, C-37(4):440–448, 1988.